

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 15:45:25 ; Search time 19.72 Seconds

(without alignments)
30.811 Million cell updates/sec

Title: US-09-609-721-90

Perfect score: 59

Sequence: 1 XXXXXXXXXXXXXCXPBCXXXXXX 27

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 segs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database : Issued Patents, AA:*

1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfilist1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	61.0	52	2	US-08-799-173A-8
2	36	61.0	59	1	US-07-862-021B-15
3	36	61.0	59	5	PCT-US93-03164-15
4	36	61.0	220	2	US-08-417-495-26
5	36	61.0	220	2	US-08-284-391B-26
6	36	61.0	220	4	US-09-218-950-26
7	36	61.0	220	5	PCT-US92-01785-26
8	36	61.0	220	5	PCT-US93-00454-26
9	36	61.0	392	2	US-08-799-173A-7
10	36	61.0	568	1	US-07-862-021B-14
11	36	61.0	568	5	PCT-US93-03164-14
12	36	61.0	802	1	US-07-862-021B-12
13	36	61.0	802	1	US-08-313-288B-12
14	36	61.0	802	5	PCT-US93-03164-12
15	36	61.0	807	1	US-07-862-021B-10
16	36	61.0	807	5	PCT-US93-03164-10
17	36	61.0	807	5	PCT-US93-03164-10
18	36	61.0	1196	1	US-08-144-121-4
19	36	61.0	1196	2	US-08-735-893-4
20	31	52.5	22	3	US-08-256-747C-1
21	31	52.5	22	4	US-08-834-130A-1
22	31	52.5	25	3	US-08-256-747C-57
23	31	52.5	25	4	US-08-834-130A-57
24	31	52.5	149	3	US-08-834-224-4
25	31	52.5	481	1	US-07-927-071-1
26	31	52.5	481	6	5164481-1
27	31	52.5	992	1	US-08-127-499A-1

28	31	52.5	992	1	US-08-482-847-1	Sequence 1, Appl
29	31	52.5	1063	1	US-08-093-453B-3	Sequence 3, Appl
30	31	52.5	1063	1	US-08-127-499A-8	Sequence 8, Appl
31	31	52.5	1063	1	US-08-482-847-8	Sequence 8, Appl
32	31	52.5	1214	2	US-08-231-193A-54	Sequence 54, Appl
33	31	52.5	1214	3	US-08-486-273A-54	Sequence 54, Appl
34	31	52.5	1214	3	US-08-480-474-54	Sequence 54, Appl
35	31	52.5	1214	3	US-08-940-086A-54	Sequence 54, Appl
36	31	52.5	1219	2	US-08-231-193A-50	Sequence 50, Appl
37	31	52.5	1219	2	US-08-486-273A-50	Sequence 50, Appl
38	31	52.5	1219	3	US-08-480-474-50	Sequence 50, Appl
39	31	52.5	1219	3	US-08-940-086A-50	Sequence 50, Appl
40	31	52.5	1231	2	US-08-231-193A-48	Sequence 48, Appl
41	31	52.5	1231	2	US-08-486-273A-48	Sequence 48, Appl
42	31	52.5	1231	3	US-08-480-474-48	Sequence 48, Appl
43	31	52.5	1231	3	US-08-940-086A-48	Sequence 48, Appl
44	31	52.5	1236	2	US-08-231-193A-6	Sequence 6, Appl
45	31	52.5	1236	2	US-08-486-273A-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-799-173A-8
Sequence 8, Application US/08799173A
Patent No. 5871969
GENERAL INFORMATION: 216199
APPLICANT: HASTINGS, GREGG,
APPLICANT: PATRICK J. DILLON
TITLE OF INVENTION: HUMAN NEURONAL ATTACHMENT FACTOR-1
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/799,173A
FILING DATE: 11-FEB-1997
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PR226
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 52 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-799-173A-8

Query Match 61.0%; Score 36; DB 2; Length 52;
Best Local Similarity 83.3%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 1; Indels 0;

OY 15 CXGPGC 20
I I I I I
Db 47 CMGPGC 52

RESULT 2
US-07-862-021B-15
; Sequence 15, Application US/07862021B
; Patent No. 5279666 1/8194
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M
; APPLICANT: Klar, Avihu
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/862,021B
; FILING DATE: 19920405
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 59 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-862-021B-15

Query Match 61.0%; Score 36; DB 1; Length 59;
Best Local Similarity 83.3%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 15 CXGPGC 20
DB 49 CMGPGC 54

RESULT 3
PCT-US93-03164-15
; Sequence 15, Application PC/TUS9303164
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M
; APPLICANT: Klar, Avihu
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/03164
FILING DATE: 19930402
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 40028
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 59 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-03164-15

Query Match 61.0%; Score 36; DB 5; Length 59;
Best Local Similarity 83.3%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 15 CXGPGC 20
DB 49 CMGPGC 54

RESULT 4
US-08-417-495-26
; Sequence 26, Application US/08417495
; Patent No. 5843728
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian et al.
; TITLE OF INVENTION: Redirection of Cellular Immunity by
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50x or 55x
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: Wordperfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/417,495
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/203,866
; FILING DATE:
; APPLICATION NUMBER: US/07/847,566
; FILING DATE:
; APPLICATION NUMBER: 07/665,961
; FILING DATE: March 7, 1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/119002

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 220 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: amino acids
US-08-417-495-26

Query Match 61.0%; Score 36; DB 2; Length 220;
Best Local Similarity 83.3%; Pred. No. 48;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 15 CXGPC 20
11111
Db 21 CLGPC 26

RESULT 5

US-08-284-391B-26
Sequence 26: Application US/08284391B
Patent No. 3851828
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391B
FILING DATE: 02-AUG-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 220 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-284-391B-26

Query Match 61.0%; Score 36; DB 2; Length 220;
Best Local Similarity 83.3%; Pred. No. 48;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 15 CXGPC 20
11111
Db 21 CLGPC 26

RESULT 6

US-09-218-950-26
Sequence 26: Application US/09218950
Patent No. 6284240
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/218,950
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391
FILING DATE: 02-AUG-1994
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 220 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-218-950-26

Query Match 61.0%; Score 36; DB 4; Length 220;
Best Local Similarity 83.3%; Pred. No. 48;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 15 CXGPC 20
11111
Db 21 CLGPC 26

RESULT 7
PCT-US92-01785-26
Sequence 26, Application PC/TUS9201785
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor
TITLE OF INVENTION: Chimera
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50z or 55sx
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/01785
FILING DATE: 19920306
CLASSIFICATION: 530
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 220 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: amino acids
PCT-US92-01785-26

Query Match 61.0%; Score 36; DB 5; Length 220;
Best Local Similarity 83.3%; Pred. No. 48;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 CXGPGC 20
| | | | |
DB 21 CLGPGC 26

RESULT 8
PCT-US95-00454-26
Sequence 26, Application PC/TUS9500454
GENERAL INFORMATION:
APPLICANT: Seed, Brian et al.
TITLE OF INVENTION: Targeted Cytolysis of HIV-Infected
TITLE OF INVENTION: Cells by Chimeric CD4 Receptor-
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM PS/2 Model 50z or 55sx
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00454
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/847,566
FILING DATE: March 6, 1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 220 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: amino acids
PCT-US95-00454-26

Query Match 61.0%; Score 36; DB 5; Length 220;
Best Local Similarity 83.3%; Pred. No. 48;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 CXGPGC 20
| | | | |
DB 21 CLGPGC 26

RESULT 9
US-08-799-173A-7
Sequence 7, Application US/08799173A
Patent No. 5871969
GENERAL INFORMATION:
APPLICANT: HASTINGS, GREGG,
APPLICANT: PATRICK J. DILLON
TITLE OF INVENTION: HUMAN NEURONAL ATTACHMENT FACTOR-1
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/799,173A
FILING DATE: 11-FEB-1997
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: BROOKS, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF226
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512

;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 392 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-799-173A-7

Query Match 61.0%; Score 36; DB 2; Length 392;
Best Local Similarity 83.3%; Pred. No. 74;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 CXGPC 20
| | | | |
Db 339 CMGPC 344

RESULT 10
US-07-862-021B-14
; Sequence 14, Application US/07862021B
; Patent No. 5279966
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M
; ATTORNEY/AGENT INFORMATION:
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/862,021B
; FILING DATE: 19920405
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 568 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-862-021B-14

Query Match 61.0%; Score 36; DB 1; Length 568;
Best Local Similarity 83.3%; Pred. No. 98;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 CXGPC 20
| | | | |
Db 248 CMGPC 253

RESULT 11
PCT-US93-03164-14

;; Sequence 14, Application PC/TUS9303164
;; GENERAL INFORMATION:
;; APPLICANT: Jessell, Thomas M
;; ATTORNEY/AGENT INFORMATION:
;; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
;; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
;; NUMBER OF SEQUENCES: 20
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Cooper & Dunham
;; STREET: 30 Rockefeller Plaza
;; CITY: New York
;; STATE: New York
;; COUNTRY: USA
;; ZIP: 10112
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: IBM PC compatible
;; SOFTWARE: Patent Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US93/03164
;; FILING DATE: 19930402
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: White, John P
;; REGISTRATION NUMBER: 28,678
;; REFERENCE/DOCKET NUMBER: 40028
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 977-9550
;; TELEFAX: (212) 664-0525
;; TELEX: 422523 COOP UI
;; INFORMATION FOR SEQ ID NO: 14:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 568 amino acids
;; TYPE: AMINO ACID
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
PCT-US93-03164-14

Query Match 61.0%; Score 36; DB 5; Length 568;
Best Local Similarity 83.3%; Pred. No. 98;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 CXGPC 20
| | | | |
Db 248 CMGPC 253

RESULT 12
US-07-862-021B-12
; Sequence 12, Application US/07862021B
; Patent No. 5279966
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M
; ATTORNEY/AGENT INFORMATION:
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/862,021B
FILING DATE: 19920405
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 40028
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 802 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-862-021B-12

Query Match 61.0%; Score 36; DB 1; Length 802;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 CXGPCC 20
11111
Db 484 CMGPCC 489

RESULT 13
US-08-313-288B-12
Sequence 12, Application US/08313288B
Patent No. 5750502
GENERAL INFORMATION:
APPLICANT: Jessell, Thomas M. and Avihu Klar
TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
City: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/313,288B
FILING DATE: January 5, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0526
TELEX:
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 802 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-313-288B-12

Query Match 61.0%; Score 36; DB 1; Length 802;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 CXGPCC 20
11111
Db 484 CMGPCC 489

RESULT 14
PCT-US93-03164-12
Sequence 12, Application PC/TUS9303164
GENERAL INFORMATION:
APPLICANT: Jessell, Thomas M
APPLICANT: Klar, Avihu
TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
City: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/03164
FILING DATE: 19930402
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 40028
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 802 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-03164-12

Query Match 61.0%; Score 36; DB 5; Length 802;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 CXGPCC 20
11111
Db 484 CMGPCC 489

RESULT 15
US-07-862-021B-10
Sequence 10, Application US/07862021B
Patent No. 5279966
GENERAL INFORMATION:
APPLICANT: Jessell, Thomas M
APPLICANT: Klar, Avihu
TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/862,021B
; FILING DATE: 19920405
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 807 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-862-021B-10

Query Match 61.0%; Score 36; DB 1; Length 807;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 15 CXGPGC 20
I I I I I
Db 489 CMGPGC 494

Search completed: November 21, 2001, 15:45:26
Job time: 29 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 15:46:04 ; Search time 32.74 Seconds

(without alignments)
49.995 Million cell updates/sec

Title: US-09-609-721-90

Perfect score: 59
Sequence: 1 xxxxxxxxxxxxxcgcgcxxxxxxx 27

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq_0601.*
1: /SID58/gcgcdata/geneseq/geneseqp/AA1980.DAT.*
2: /SID58/gcgcdata/geneseq/geneseqp/AA1981.DAT.*
3: /SID58/gcgcdata/geneseq/geneseqp/AA1982.DAT.*
4: /SID58/gcgcdata/geneseq/geneseqp/AA1983.DAT.*
5: /SID58/gcgcdata/geneseq/geneseqp/AA1984.DAT.*
6: /SID58/gcgcdata/geneseq/geneseqp/AA1985.DAT.*
7: /SID58/gcgcdata/geneseq/geneseqp/AA1986.DAT.*
8: /SID58/gcgcdata/geneseq/geneseqp/AA1987.DAT.*
9: /SID58/gcgcdata/geneseq/geneseqp/AA1988.DAT.*
10: /SID58/gcgcdata/geneseq/geneseqp/AA1989.DAT.*
11: /SID58/gcgcdata/geneseq/geneseqp/AA1990.DAT.*
12: /SID58/gcgcdata/geneseq/geneseqp/AA1991.DAT.*
13: /SID58/gcgcdata/geneseq/geneseqp/AA1992.DAT.*
14: /SID58/gcgcdata/geneseq/geneseqp/AA1993.DAT.*
15: /SID58/gcgcdata/geneseq/geneseqp/AA1994.DAT.*
16: /SID58/gcgcdata/geneseq/geneseqp/AA1995.DAT.*
17: /SID58/gcgcdata/geneseq/geneseqp/AA1996.DAT.*
18: /SID58/gcgcdata/geneseq/geneseqp/AA1997.DAT.*
19: /SID58/gcgcdata/geneseq/geneseqp/AA1998.DAT.*
20: /SID58/gcgcdata/geneseq/geneseqp/AA1999.DAT.*
21: /SID58/gcgcdata/geneseq/geneseqp/AA2000.DAT.*
22: /SID58/gcgcdata/geneseq/geneseqp/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	62.7	20	22	AA196413
2	37	62.7	161	20	AA1921843
3	37	62.7	204	20	AA1913498
4	36	61.0	6	22	AA196420
5	36	61.0	16	22	AA196381
6	36	61.0	16	22	AA196382
7	36	61.0	16	22	AA196383
8	36	61.0	16	22	AA196384
9	36	61.0	16	22	AA196385
10	36	61.0	16	22	AA196386
11	36	61.0	16	22	AA196387

12	36	61.0	16	22	AA196388	ErbB2 binding pept
13	36	61.0	16	22	AA196390	ErbB2 binding pept
14	36	61.0	16	22	AA196391	ErbB2 binding pept
15	36	61.0	16	22	AA196392	ErbB2 binding pept
16	36	61.0	17	22	AA196393	ErbB2 binding pept
17	36	61.0	17	22	AA196394	ErbB2 binding pept
18	36	61.0	17	22	AA196395	ErbB2 binding pept
19	36	61.0	17	22	AA196396	ErbB2 binding pept
20	36	61.0	17	22	AA196397	ErbB2 binding pept
21	36	61.0	17	22	AA196398	ErbB2 binding pept
22	36	61.0	17	22	AA196399	ErbB2 binding pept
23	36	61.0	17	22	AA196400	ErbB2 binding pept
24	36	61.0	17	22	AA196401	ErbB2 binding pept
25	36	61.0	17	22	AA196402	ErbB2 binding pept
26	36	61.0	17	22	AA196403	ErbB2 binding pept
27	36	61.0	17	22	AA196404	ErbB2 binding pept
28	36	61.0	17	22	AA196405	ErbB2 binding pept
29	36	61.0	17	22	AA196406	ErbB2 binding pept
30	36	61.0	19	22	AA196407	ErbB2 binding pept
31	36	61.0	20	22	AA196408	ErbB2 binding pept
32	36	61.0	20	22	AA196409	ErbB2 binding pept
33	36	61.0	20	22	AA196410	ErbB2 binding pept
34	36	61.0	20	22	AA196411	ErbB2 binding pept
35	36	61.0	20	22	AA196412	ErbB2 binding pept
36	36	61.0	20	22	AA196413	ErbB2 binding pept
37	36	61.0	20	22	AA196414	ErbB2 binding pept
38	36	61.0	20	22	AA196415	ErbB2 binding pept
39	36	61.0	20	22	AA196416	ErbB2 binding pept
40	36	61.0	20	22	AA196417	ErbB2 binding pept
41	36	61.0	20	22	AA196418	ErbB2 binding pept
42	36	61.0	20	22	AA196419	ErbB2 binding pept
43	36	61.0	20	22	AA196420	ErbB2 binding pept
44	36	61.0	20	22	AA196421	ErbB2 binding pept
45	36	61.0	20	22	AA196422	ErbB2 binding pept

ALIGNMENTS

RESULT 1
ID AAB76413 standard; Peptide: 20 AA.
AC AAB76413:
DF 10-APR-2001 (first entry)
XX
XX
DE ErbB2 binding peptide amino acid sequence SEQ ID 64.
XX
XX Human; erbB2; HER2; cancer; nervous system disease; stroke; ischaemia;
KW metabolic disorder; nutritional deficiency; Alzheimer's disease;
KW Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
KW Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.
XX
XX Synthetic.
OS
OS
PN WO200101748-A2.
XX
XX 11-JAN-2001.
XX
XX 30-JUN-2000; 2000WO-US18283.
XX
XX 02-JUL-1999; 99US-0142232.
XX
XX (GETH) GENENTECH INC.
XX
XX Dennis MS;
XX WPI; 2001-123048/13.
XX
XX Non-naturally occurring peptide ligands which compete for binding human
XX erbB2 gene products, useful for treating e.g. Alzheimer's disease,
XX multiple sclerosis and diabetic neuropathy -

XX PS Disclosure; Figure 15; 116pp; English.

CC CC This invention relates to non-naturally occurring peptide ligands which bind to the human erbB2 gene product ErbB2 (also known as HER2). Peptides represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples of the ErbB2 binding ligands of the invention. Sequences

CC AAB76421 - AAB76431 represent antibody Fc amino acid sequences used in the isolation of the peptides of the invention. The peptides compete for binding ErbB2 with naturally occurring ligands, and may be used to treat disorders characterized by over expression of HER2/ErbB2 such as cancers, diseases of the nervous system, musculature and epithelia, e.g. nervous system damage resulting from trauma, surgery, strokes, ischaemia, infection, metabolic disorders, nutritional deficiency or toxic agents.

CC In particular the synthetic peptide ligands may be used to treat Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's chorea, Down's syndrome, nerve deafness, Meniere's disease and neuropathy associated with diabetes.

CC XX

SQ Sequence 20 AA;

QY 15 CXGPGC 20
I I I I I

Db 8 ctgpgc 13

Query Match 62.7%; Score 37; DB 22; Length 20;
Best Local Similarity 83.3%; Pred. No. 15;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 2
AAY21843
ID AAY21843 standard; Protein; 161 AA.

AC AAY21843;
XX

DT 20-SEP-1999 (first entry)

DE Human signal peptide-containing protein (SIGP) (clone ID 1534876).

XX

KW Signal-peptide containing protein; SIGP; human; cancer; immune response; adenocarcinoma; leukemia; lymphoma; melanoma; myeloma; sarcoma; AIDS; Addison's disease; adult respiratory distress syndrome; allergy; anemia; asthma; atherosclerosis; bronchitis; cholecystitis; Crohn's disease; ulcerative colitis; atopic dermatitis; dermatomyositis; emphysema; diabetes mellitus; atrophic gastritis; glomerulonephritis; gout; trauma; Grave's Disease; hyper eosinophilia; irritable bowel syndrome; infection; lupus erythematosus; multiple sclerosis; myasthenia gravis; inflammation; osteoarthritis; osteoporosis; pancreatitis; polymyositis; scleroderma; rheumatoid arthritis; Sjogren's syndrome; autoimmune thyroiditis.

XX

OS Homo sapiens.

XX

PN WO9933981-A2.

XX

PD 08-JUL-1999.

XX

PF 22-DEC-1998; 98WO-US27598.

XX

PR 31-DEC-1997; 97US-0002485. 09/002485

XX

PA (INCY-) INCYTE PHARM INC.

XX

PI Baughn MR, Corley NC, Guegler KJ, Hillman JL, Lal P;

PI Sather SK, Shah P;

XX

DR WPI; 1999-430242/36.

XX

DR N-PSDB; AAX82078.

XX

PT Human signal-peptide containing protein coding sequences used to treat cancer and immune responses

XX

PS Claim 1; Page 78; 99pp; English.

XX

CC The invention provides human signal-peptide containing proteins (SIGP) (AAY21841-855) and polynucleotides (AAX82076-90) encoding the proteins.

CC A host cell containing a vector comprising SIGP DNA can be used to produce the SIGP protein. The SIGP protein can be used, in conjunction with a pharmaceutical carrier to treat or prevent a cancer. An antagonist of the SIGP protein can be used to treat or prevent a cancer or an immune response. The cancers that can be treated or prevented include sarcomas, adenocarcinomas, leukemias, lymphomas, melanomas, teratocarcinomas, myelomas and cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus. The immune responses that can be treated or prevented include, AIDS, Addison's disease, adult respiratory distress syndrome, allergies, anemia, asthma, atherosclerosis, bronchitis, cholecystitis, Crohn's disease, ulcerative colitis, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, atrophic gastritis, glomerulonephritis, Grave's disease, gout, hyper eosinophilia, irritable bowel syndrome, lupus erythematosus, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, rheumatoid arthritis, scleroderma, Sjogren's syndrome, and autoimmune thyroiditis, complications of cancer, infections, and trauma.

CC XX

SQ Sequence 161 AA;

QY 15 CXGPGC 20
I I I I I

Db 6 ctgpgc 11

Query Match 62.7%; Score 37; DB 20; Length 161;
Best Local Similarity 83.3%; Pred. No. 61;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 3
AAY13498
ID AAY13498 standard; Protein; 204 AA.

AC AAY13498;
XX

DT 30-JUL-1999 (first entry)

DE Tissue cement protein fragment encoded by clone 24.

XX

KW Tissue cement protein; TCP; blood-feeding; ectoparasite; therapy; pharmaceutical; vaccine; tissue bonding; surgical wound; laceration; skin grafting; immunogen; infection; arthropod parasite.

XX

OS Rhinipcephalus appendiculatus.

XX

PN WO9924567-A1.

XX

PD 20-MAY-1999.

XX

PF 12-NOV-1998; 98WO-GB03397.

XX

PR 12-NOV-1997; 97GB-0023945.

XX

PA (NAEN-) NAT ENVIRONMENTAL RES COUNCIL.

XX

PI Nuttall Patricia A, Paesen GC;

PI

DR WPI; 1999-327399/27.

XX

DR N-PSDB; AAX55697.

XX

PT Tissue cement proteins produced by blood-feeding ectoparasites and related polynucleotides

XX

PS Disclosure; Fig 5; 65pp; English.

XX The invention relates to tissue cement proteins (TCPs) produced by
CC blood-feeding ectoparasites. The TCP, in a pharmaceutical composition,
CC is useful for therapy, as a vaccine or vaccine component. The TCP itself
CC is used to immunize an animal for production of such a vaccine. The TCPs
CC are useful for bonding animal tissues. This may be temporary, or permanent
CC bonding and used in, e.g. repair of lacerated surgical wounds, lacerations, in
CC skin grafting, etc. The TCP can also be used as a protective immunogen in
CC the control of diseases caused by infections transmitted by arthropod
CC parasites. The tick TCPs provide a non-immunogenic tissue cement capable
CC of bonding mammalian tissue with great strength. The hardening or elastic
CC properties of the TCPs can be tailored for particular requirements.

XX Sequence 204 AA:

Query Match 62.7%; Score 37; DB 20; Length 204;
Best Local Similarity 83.3%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 CXGPGC 20
I I I I I
Db 44 cagpgc 49

RESULT 4

AAB76420
ID AAB76420 standard; Peptide; 6 AA.

AC AAB76420;

DT 10-APR-2001 (first entry)

DE Erbb2 binding peptide amino acid sequence SEQ ID 71.

XX Human; erbb2; HER2; cancer; nervous system disease; stroke; ischemia;
KM metabolic disorder; nutritional deficiency; Alzheimer's disease;
KM Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
KM Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.
XX Synthetic.

PN WO200101748-A2.

PD 11-JAN-2001.

PF 30-JUN-2000; 2000WO-US18283.

PR 02-JUL-1999; 99US-0142232.

PA (GETH) GENENTECH INC.

PI Dennis MS;

DR WPI; 2001-123048/13.

PT Non-naturally occurring peptide ligands which compete for binding human
erB2 gene products, useful for treating e.g. Alzheimer's disease,
PT multiple sclerosis and diabetic neuropathy -

PS Disclosure; Page 3; 116pp; English.

XX This invention relates to non-naturally occurring peptide ligands which
CC bind to the human erbb2 gene product Erbb2 (also known as HER2). Peptides
CC represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples
CC of the Erbb2 binding ligands of the invention. Sequences
CC AAB76421 - AAB76431 represent antibody Fc amino acid sequences used in
CC the isolation of the peptides of the invention. The peptides compete for
CC binding Erbb2 with naturally occurring ligands, and may be used to treat
CC disorders characterized by over expression of HER2/Erbb2 such as cancers,
CC diseases of the nervous system, musculature and epithelia, e.g. nervous
CC system damage resulting from trauma, surgery, strokes, ischemia,
CC infection, metabolic disorders, nutritional deficiency or toxic agents.

CC In particular the synthetic peptide ligands may be used to treat
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis,
CC Huntington's chorea, Down's syndrome, nerve deafness, Meniere's disease
CC and neuropathy associated with diabetes.

XX Sequence 6 AA:

Query Match 61.0%; Score 36; DB 22; Length 6;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 CXGPGC 20
I I I I I
Db 1 cxgpgc 6

RESULT 5

AAB76381
ID AAB76381 standard; Peptide; 16 AA.

AC AAB76381;

DT 10-APR-2001 (first entry)

DE Erbb2 binding peptide amino acid sequence SEQ ID 32.

XX Human; erbb2; HER2; cancer; nervous system disease; stroke; ischemia;
KM metabolic disorder; nutritional deficiency; Alzheimer's disease;
KM Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
KM Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.

XX Synthetic.

PN WO200101748-A2.

PD 11-JAN-2001.

PF 30-JUN-2000; 2000WO-US18283.

PR 02-JUL-1999; 99US-0142232.

PA (GETH) GENENTECH INC.

PI Dennis MS;

DR WPI; 2001-123048/13.

PT Non-naturally occurring peptide ligands which compete for binding human
erB2 gene products, useful for treating e.g. Alzheimer's disease,
PT multiple sclerosis and diabetic neuropathy -

PS Disclosure; Figure 16; 116pp; English.

XX This invention relates to non-naturally occurring peptide ligands which
CC bind to the human erbb2 gene product Erbb2 (also known as HER2). Peptides
CC represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples
CC of the Erbb2 binding ligands of the invention. Sequences
CC AAB76421 - AAB76431 represent antibody Fc amino acid sequences used in
CC the isolation of the peptides of the invention. The peptides compete for
CC binding Erbb2 with naturally occurring ligands, and may be used to treat
CC disorders characterized by over expression of HER2/Erbb2 such as cancers,
CC diseases of the nervous system, musculature and epithelia, e.g. nervous
CC system damage resulting from trauma, surgery, strokes, ischemia,
CC infection, metabolic disorders, nutritional deficiency or toxic agents.
CC In particular the synthetic peptide ligands may be used to treat
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis,
CC Huntington's chorea, Down's syndrome, nerve deafness, Meniere's disease
CC and neuropathy associated with diabetes.

XX Sequence 16 AA:

Query Match 61.0%; Score 36; DB 22; Length 16;
Best Local Similarity 83.3%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 CXGPGC 20
I I I I I
Db 8 c1gpgc 13

RESULT 6
AAB76382
ID AAB76382 standard; Peptide: 16 AA.
XX
AC AAB76382;
XX
DT 10-APR-2001 (first entry)
XX
DE ErbB2 binding peptide amino acid sequence SEQ ID 33.
XX
KW Human; erbB2; HER2; cancer; nervous system disease; stroke; ischaemia;
KW metabolic disorder; nutritional deficiency; Alzheimer's disease;
KW Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
KW Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.
XX
OS Synthetic.
XX
PN WO200101748-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18283.
XX
PR 02-JUL-1999; 99US-0142232.
XX
PA (GETH) GENENTECH INC.
XX
PI Dennis MS;
XX
DR WPI; 2001-123048/13.
XX
PT Non-naturally occurring peptide ligands which compete for binding human
PT erb2 gene products, useful for treating e.g. Alzheimer's disease,
PT multiple sclerosis and diabetic neuropathy -
XX
PS Disclosure; Figure 16; 116pp; English.
XX
CC This invention relates to non-naturally occurring peptide ligands which
CC bind to the human erbB2 gene product ErbB2 (also known as HER2). Peptides
CC represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples
CC of the ErbB2 binding ligands of the invention. Sequences
CC AAB76421 - AAB76431 represent antibody Fc amino acid sequences used in
CC the isolation of the peptides of the invention. The peptides compete for
CC binding ErbB2 with naturally occurring ligands, and may be used to treat
CC disorders characterized by over expression of HER2/ErbB2 such as cancers,
CC diseases of the nervous system, musculature and epithelia, e.g. nervous
CC system damage resulting from trauma, surgery, strokes, ischaemia,
CC infection, metabolic disorders, nutritional deficiency or toxic agents.
CC In particular the synthetic peptide ligands may be used to treat
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis,
CC Huntington's chorea, Down's syndrome, nerve deafness, Meniere's disease
CC and neuropathy associated with diabetes.
XX
SQ Sequence 16 AA;

Query Match 61.0%; Score 36; DB 22; Length 16;
Best Local Similarity 83.3%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 CXGPGC 20
I I I I I
Db 8 c1gpgc 13

RESULT 7
AAB76383
ID AAB76383 standard; Peptide: 16 AA.
XX
AC AAB76383;
XX
DT 10-APR-2001 (first entry)
XX
DE ErbB2 binding peptide amino acid sequence SEQ ID 34.
XX
KW Human; erbB2; HER2; cancer; nervous system disease; stroke; ischaemia;
KW metabolic disorder; nutritional deficiency; Alzheimer's disease;
KW Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
KW Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.
XX
OS Synthetic.
XX
PN WO200101748-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18283.
XX
PR 02-JUL-1999; 99US-0142232.
XX
PA (GETH) GENENTECH INC.
XX
PI Dennis MS;
XX
DR WPI; 2001-123048/13.
XX
PT Non-naturally occurring peptide ligands which compete for binding human
PT erb2 gene products, useful for treating e.g. Alzheimer's disease,
PT multiple sclerosis and diabetic neuropathy -
XX
PS Disclosure; Figure 16; 116pp; English.
XX
CC This invention relates to non-naturally occurring peptide ligands which
CC bind to the human erbB2 gene product ErbB2 (also known as HER2). Peptides
CC represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples
CC of the ErbB2 binding ligands of the invention. Sequences
CC AAB76421 - AAB76431 represent antibody Fc amino acid sequences used in
CC the isolation of the peptides of the invention. The peptides compete for
CC binding ErbB2 with naturally occurring ligands, and may be used to treat
CC disorders characterized by over expression of HER2/ErbB2 such as cancers,
CC diseases of the nervous system, musculature and epithelia, e.g. nervous
CC system damage resulting from trauma, surgery, strokes, ischaemia,
CC infection, metabolic disorders, nutritional deficiency or toxic agents.
CC In particular the synthetic peptide ligands may be used to treat
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis,
CC Huntington's chorea, Down's syndrome, nerve deafness, Meniere's disease
CC and neuropathy associated with diabetes.
XX
SQ Sequence 16 AA;

Query Match 61.0%; Score 36; DB 22; Length 16;
Best Local Similarity 83.3%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 CXGPGC 20
I I I I I
Db 8 c1gpgc 13

RESULT 8
AAB76384
ID AAB76384 standard; Peptide: 16 AA.
XX
AC AAB76384;
XX
DT 10-APR-2001 (first entry)

XX	Erbb2 binding peptide amino acid sequence SEQ ID 35.
DE	
XX	Human; erbb2; HER2; cancer; nervous system disease; stroke; ischaemia;
XX	metabolic disorder; nutritional deficiency; Alzheimer's disease;
KW	Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
KW	Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.
XX	
OS	Synthetic.
XX	
PN	WO200101748-A2.
XX	
XX	11-JAN-2001.
XX	
XX	30-JUN-2000; 2000WO-US18283.
XX	
XX	02-JUL-1999; 99US-0142232.
XX	
PA	(GETH) GENENTECH INC.
XX	
P1	Dennis MS;
DR	WPI; 2001-123048/13.
XX	
XX	Non-naturally occurring peptide ligands which compete for binding human
PT	erx2 gene products, useful for treating e.g. Alzheimer's disease,
PT	multiple sclerosis and diabetic neuropathy -
PS	Disclosure: Figure 16; 116pp; English.
XX	
XX	This invention relates to non-naturally occurring peptide ligands which
CC	bind to the human erbb2 gene product Erbb2 (also known as HER2). Peptides
CC	represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples
CC	of the Erbb2 binding ligands of the invention. Sequences
CC	AAB76421 - AAB76431 represent antibody Fc amino acid sequences used in
CC	the isolation of the peptides of the invention. The peptides compete for
CC	binding Erbb2 with naturally occurring ligands, and may be used to treat
CC	disorders characterized by over expression of HER2/Erbb2 such as cancers,
CC	diseases of the nervous system, musculature and epithelia, e.g. nervous
CC	system damage resulting from trauma, surgery, strokes, ischaemia,
CC	infection, metabolic disorders, nutritional deficiency or toxic agents.
CC	In particular the synthetic peptide ligands may be used to treat
CC	Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis,
CC	Huntington's chorea, Down's syndrome, nerve deafness, Meniere's disease
CC	and neuropathy associated with diabetes.
SO	Sequence 16 AA:
XX	
XX	Query Match 61.0%; Score 36; DB 22; Length 16;
XX	Best Local Similarity 83.3%; Pred. NO. 18;
XX	Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY	15 CXGPGC 20
DB	1 1111
DB	8 c1gpgc 13
XX	
XX	RESULT 9
XX	AAB76385
XX	ID AAB76385 standard; Peptide; 16 AA.
XX	
XX	AAB76385;
XX	
XX	10-APR-2001 (first entry)
XX	
DE	Erbb2 binding peptide amino acid sequence SEQ ID 36.
XX	
XX	Human; erbb2; HER2; cancer; nervous system disease; stroke; ischaemia;
KW	metabolic disorder; nutritional deficiency; Alzheimer's disease;
KW	Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
KW	Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.
XX	

OS	Synthetic.
XX	
PN	WO200101748-A2.
XX	
PD	11-JAN-2001.
XX	
Pf	30-JUN-2000; 2000WO-US18283.
XX	
PR	02-JUL-1999; 99US-0142232.
XX	
PA	(GETH) GENENTECH INC.
XX	
PI	Dennis MS;
XX	
DR	WPI; 2001-123048/13.
XX	
PT	Non-naturally occurring peptide ligands which compete for binding human
PT	erbB gene products, useful for treating e.g. Alzheimer's disease,
PT	multiple sclerosis and diabetic neuropathy -
XX	
PS	Disclosure; Figure 16; 11pp; English.
XX	
CC	This invention relates to non-naturally occurring peptide ligands which
CC	bind to the human erbB2 gene product ErbB2 (also known as HER2). Peptides
CC	represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples
CC	of the ErbB2 binding ligands of the invention. Sequences
CC	AAB76421 - AAB76431 represent antibody Fc amino acid sequences used in
CC	the isolation of the peptides of the invention. The peptides compete for
CC	binding ErbB2 with naturally occurring ligands, and may be used to treat
CC	disorders characterized by over expression of HER2/ErbB2 such as cancers,
CC	diseases of the nervous system, musculature and epithelia, e.g. nervous
CC	system damage resulting from trauma, surgery, strokes, ischaemia,
CC	infection, metabolic disorders, nutritional deficiency or toxic agents.
CC	In particular the synthetic peptide ligands may be used to treat
CC	Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis,
CC	Huntington's chorea, Down's syndrome, nerve deafness, Meniere's disease
CC	and neuropathy associated with diabetes.
XX	
XX	Sequence 16 AA:
QY	15 CXGPC 20
	I I I I I
DB	8 c1pgc 13
RESULT 10	
AAB76386	ID AAB76386 standard; Peptide; 16 AA.
XX	
AC	AAB76386;
XX	
DT	10-APR-2001 (first entry)
XX	
DE	ErbB2 binding peptide amino acid sequence SEQ ID 37.
XX	
KM	Human; erbB2; HER2; cancer; nervous system disease; stroke; ischaemia;
KM	metabolic disorder; nutritional deficiency; Alzheimer's disease;
KM	Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
KM	Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.
XX	
OS	Synthetic.
XX	
PN	WO200101748-AZ.
XX	
PD	11-JAN-2001.
XX	
Pf	30-JUN-2000; 2000WO-US18283.
XX	

PR 02-JUL-1999; 99US-0142232.
XX (GETH) GENENTECH INC.
XX
XX
PI Dennis MS;
DR WPI; 2001-123048/13.
XX
XX Non-naturally occurring peptide ligands which compete for binding human
PT erb2 gene products, useful for treating e.g. Alzheimer's disease,
PT multiple sclerosis and diabetic neuropathy -
PS Disclosure; Figure 16; 116pp; English.
XX
XX This invention relates to non-naturally occurring peptide ligands which
CC bind to the human erbB2 gene product ErbB2 (also known as HER2). Peptides
CC represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples
CC of the ErbB2 binding ligands of the invention. Sequences
CC AAB76421 - AAB76431 represent antibody Fc amino acid sequences used in
CC the isolation of the peptides of the invention. The peptides compete for
CC binding ErbB2 with naturally occurring ligands, and may be used to treat
CC disorders characterized by over expression of HER2/ErbB2 such as cancers,
CC diseases of the nervous system, musculature and epithelia, e.g. nervous
CC system damage resulting from trauma, surgery, strokes, ischaemia,
CC infection, metabolic disorders, nutritional deficiency or toxic agents.
CC In particular the synthetic peptide ligands may be used to treat
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis,
CC Huntington's chorea, Down's syndrome, nerve deafness, Meniere's disease
CC and neuropathy associated with diabetes.
XX
XX Sequence 16 AA;
SQ

Query Match 61.0%; Score 36; DB 22; Length 16;
Best Local Similarity 83.3%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 15 CXGPGC 20
I I I I I
Db 8 c1gpgc 13

RESULT 11
AAB76387
ID AAB76387 standard; Peptide; 16 AA.
XX
AC AAB76387;
XX
DT 10-APR-2001 (first entry)
XX
DE ErbB2 binding peptide amino acid sequence SEQ ID 38.
XX
XX Human; erbB2; HER2; cancer; nervous system disease; stroke; ischaemia;
KW metabolic disorder; nutritional deficiency; Alzheimer's disease;
KW Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
KW Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.
XX
OS Synthetic.
XX
PN WO200101748-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18283.
XX
PR 02-JUL-1999; 99US-0142232.
XX
PA (GETH) GENENTECH INC.
XX
PI Dennis MS;
XX
DR WPI; 2001-123048/13.
XX

PT Non-naturally occurring peptide ligands which compete for binding human
PT erb2 gene products, useful for treating e.g. Alzheimer's disease,
PT multiple sclerosis and diabetic neuropathy -
PS Disclosure; Figure 16; 116pp; English.
XX
XX This invention relates to non-naturally occurring peptide ligands which
CC bind to the human erbB2 gene product ErbB2 (also known as HER2). Peptides
CC represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples
CC of the ErbB2 binding ligands of the invention. Sequences
CC AAB76421 - AAB76431 represent antibody Fc amino acid sequences used in
CC the isolation of the peptides of the invention. The peptides compete for
CC binding ErbB2 with naturally occurring ligands, and may be used to treat
CC disorders characterized by over expression of HER2/ErbB2 such as cancers,
CC diseases of the nervous system, musculature and epithelia, e.g. nervous
CC system damage resulting from trauma, surgery, strokes, ischaemia,
CC infection, metabolic disorders, nutritional deficiency or toxic agents.
CC In particular the synthetic peptide ligands may be used to treat
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis,
CC Huntington's chorea, Down's syndrome, nerve deafness, Meniere's disease
CC and neuropathy associated with diabetes.
XX
XX Sequence 16 AA;
SQ

Query Match 61.0%; Score 36; DB 22; Length 16;
Best Local Similarity 83.3%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 15 CXGPGC 20
I I I I I
Db 8 c1gpgc 13

RESULT 12
AAB76388
ID AAB76388 standard; Peptide; 16 AA.
XX
AC AAB76388;
XX
DT 10-APR-2001 (first entry)
XX
DE ErbB2 binding peptide amino acid sequence SEQ ID 39.
XX
XX Human; erbB2; HER2; cancer; nervous system disease; stroke; ischaemia;
KW metabolic disorder; nutritional deficiency; Alzheimer's disease;
KW Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
KW Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.
XX
OS Synthetic.
XX
PN WO200101748-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18283.
XX
PR 02-JUL-1999; 99US-0142232.
XX
PA (GETH) GENENTECH INC.
XX
PI Dennis MS;
XX
DR WPI; 2001-123048/13.
XX
XX Non-naturally occurring peptide ligands which compete for binding human
PT erb2 gene products, useful for treating e.g. Alzheimer's disease,
PT multiple sclerosis and diabetic neuropathy -
PS Disclosure; Figure 16; 116pp; English.
XX
XX This invention relates to non-naturally occurring peptide ligands which
CC bind to the human erbB2 gene product ErbB2 (also known as HER2). Peptides

CC represented in AAB76550 - AAB76420 and AAB76432. AAB76509 are examples
CC of the ErbB2 binding ligands of the invention. Sequences
CC AAB76421 - AAB76431 represent antibody Fc amino acid sequences used in
CC the isolation of the peptides of the invention. The peptides compete for
CC binding ErbB2 with naturally occurring ligands, and may be used to treat
CC disorders characterized by over expression of HER2/ErbB2 such as cancers
CC diseases of the nervous system, musculature and epithelia, e.g. nervous
CC system damage resulting from trauma, surgery, strokes, ischaemia,
CC infection, metabolic disorders, nutritional deficiency or toxic agents.
CC In particular the synthetic peptide ligands may be used to treat
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis,
CC Huntington's chorea, Down's syndrome, nerve deafness, Meniere's disease
CC and neuropathy associated with diabetes.

SQ Sequence 16 AA;

Query Match	61.08;	Score 36;	DB 22;	Length 16;
Best Local Similarity	83.38;	Pred. No. 18;		
Matches	5;	Conservative	0;	Mismatches 1;
				Indels 0;
				Gaps 0

QY	15	CXGPGC	20
Db	8	cigpgc	13

RESULT 13
AAB76390

ID AAB76390 standard; Peptide; 16 AA.

AC AAB76390;

DT 10-APR-2001 (first entry)

DE Erbb2 binding peptide amino acid sequence SEQ ID 41.

KM Human; *erbb2*; HRR2; cancer; nervous system disease; stroke; ischemia;
 KM metabolic disorder; nutritional deficiency; Alzheimer's disease;
 KM Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
 KM Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy

OS Synthetic.

PN WO200101748-A2.

PD 11-JAN-2001.

PF 30-JUN-2000; 2000WO-US18283.

PR 02-JUL-1999; 99US-0142232.

PA (GETH) GENENTECH INC.

PI C Dennis MS;

DR WPI; 2001-123048/13.

PT	Non-naturally occurring peptide ligands which compete for binding human
PT	erbB gene products, useful for treating e.g. Alzheimer's disease,
PT	multiple sclerosis and diabetic neuropathy -

PS Disclosure; Figure 16; 116pp; English.

This invention relates to non-naturally occurring peptide ligands which bind to the human ErbB2 gene product ErbB2 (also known as HER2). Peptides represented in AAB/6350 - AAB/6420 and AAB/6432 - AAB/6509 are examples of the ErbB2 binding ligands of the invention. Sequences AAB/6421 - AAB/6431 represent antibody Fc amino acid sequences used in the isolation of the peptides of the invention. The peptides compete for the binding of ErbB2 with naturally occurring ligands, and may be used to treat disorders characterized by over expression of HER2/ErbB2 such as cancers, diseases of the nervous system, muscular and epithelia, e.g. nervous system damage resulting from trauma, surgery, strokes, ischaemia, CC

CC infection, metabolic disorders, nutritional deficiency or toxic agents.
CC In particular the synthetic peptide ligands may be used to treat
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis,
CC Huntington's chorea, Down's syndrome, nerve deafness, Meniere's disease
CC and neuropathy associated with diabetes.

SQ Sequence 16 AA;

Query Match	61.0%	Score 36	DB 22	Length 16
Best Local Similarity	83.3%	Pred. No. 18		
Matches 5; Conservative	0	Mismatches 1	Indels 0	Gaps 0

QY	15	CXGPGC	20
Db	8	ci:pgpc	13

RESULT 14
AAR76391

ID	Peptide	AA
AAB76391	standard; Peptide; 16	AA.

AC AAB76391;

DT 10-APR-2001 (first entry)

DE	ErbB2 binding peptide amino acid sequence SEQ ID 42.
DE	ERBB2 binding peptide amino acid sequence SEQ ID 42.

KM Human; erbB2; HER2; cancer; nervous system disease; stroke; ischemia;
 KM metabolic disorder; nutritional deficiency; Alzheimer's disease;
 KM Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
 KM Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy

OS Synthetic.

PN W0200101748-A2.

PD 11-JAN-2001.

PF 30-JUN-2000; 2000WO-US18283.

PR 02-JUL-1999; 99US-0142232.

PA (GETH) GENENTECH INC.

PI Dennis MS;

DR WPI; 2001-123048/13.

PT Non-naturally occurring peptide ligands which compete for binding human
PT erb2 gene products, useful for treating e.g. Alzheimer's disease,
PT multiple sclerosis and diabetic neuropathy -

PS Disclosure; Figure 16; 116pp; English.

CC This invention relates to non-naturally occurring peptide ligands which
CC bind to the human ErbB2 gene product ErbB2 (also known as HER2). Peptides
CC represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples
CC of the ErbB2 binding ligands of the invention. Sequences
CC AAB76421 - AAB76431 represent antibody Fc amino acid sequences used in
CC the isolation of the peptides of the invention. The peptides compete for
CC binding ErbB2 with naturally occurring ligands, and may be used to treat
CC disorders characterized by over expression of HER2/erbB2 such as cancers,
CC diseases of the nervous system, musculature and epithelia, e.g. nervous
CC system damage resulting from trauma, surgery, strokes, ischaemia,
CC infection, metabolic disorders, nutritional deficiency or toxic agents.
CC In particular the synthetic peptide ligands may be used to treat
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis,
CC Huntington's chorea, Down's syndrome, nerve deafness, Meniere's disease
CC and neuropathy associated with diabetes.

... Sequence 16 AA;
SQ

Query Match 61.0%; Score 36; DB 22; Length 16;
Best Local Similarity 83.3%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Search completed: November 21, 2001, 15:46:04
Job time: 67 sec

QY 15 CXGPGC 20
| | | | |
Db 8 c1gpgc 13

RESULT 15

AAB76392
ID AAB76392 standard; Peptide; 16 AA.

XX
AC AAB76392;

XX
DT 10-APR-2001 (first entry)

XX
DE ErbB2 binding peptide amino acid sequence SEQ ID 43.

XX
KW Human: erbB2; HER2; cancer; nervous system disease; stroke; ischaemia;

KW metabolic disorder; nutritional deficiency; Alzheimer's disease;

KW Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;

KW Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.

XX
OS Synthetic.

XX
PN WO200101748-A2.

XX
PD -11-JAN-2001.

XX
PF 30-JUN-2000; 2000WO-0518283.

XX
PR 02-JUL-1999; 9905-0142232.

XX
PA (GETH) GENENTECH INC.

XX
PI Dennis MS;

XX
DR WPI; 2001-123048/13.

XX
PT Non-naturally occurring peptide ligands which compete for binding human
PT erbB2 gene products; useful for treating e.g. Alzheimer's disease,
PT multiple sclerosis and diabetic neuropathy -

XX
PS Disclosure; Figure 16; 116pp; English.

XX
CC This invention relates to non-naturally occurring peptide ligands which
CC bind to the human erbB2 gene product ErbB2 (also known as HER2). Peptides
CC represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples
CC of the ErbB2 binding ligands of the invention. Sequences
CC AAB76421 - AAB76431 represent antibody Fc amino acid sequences used in
CC the isolation of the peptides of the invention. The peptides compete for
CC binding ErbB2 with naturally occurring ligands, and may be used to treat
CC disorders characterized by over expression of HER2/ErbB2 such as cancers,
CC diseases of the nervous system, musculature and epithelia, e.g. nervous
CC system damage resulting from trauma, surgery, strokes, ischaemia,
CC infection, metabolic disorders, nutritional deficiency or toxic agents.
CC In particular the synthetic peptide ligands may be used to treat
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis,
CC Huntington's chorea, Down's syndrome, nerve deafness, Meniere's disease
CC and neuropathy associated with diabetes.

XX
SQ Sequence 16 AA;

Query Match

Best Local Similarity 61.0%; Score 36; DB 22; Length 16;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 15 CXGPGC 20
| | | | |

Db 8 c1gpgc 13

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 15:47:34 ; Search time 33.25 Seconds
(without alignments)
107.436 Million cell updates/sec

Title: US-09-609-721-90

Perfect score: 59
Sequence: 1 XXXXXXXXXXXXXCXCPCGXXXXXX 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPREMBL_16:*
1: SP_archaea:*
2: SP_bacteria:*
3: SP_fungi:*
4: SP_human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_unclassified:*
13: SP_vertebrate:*
14: SP_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	62.7	204	5	027392 trypanosoma
2	37	62.7	475	2	P71952 mycobacteri
3	36	61.0	562	4	09NUM8
4	36	61.0	614	11	09WV00
5	36	61.0	624	4	094862
6	36	61.0	656	4	09P004
7	36	61.0	656	4	09P2V7
8	36	61.0	802	13	09W770
9	36	61.0	803	13	042114
10	36	61.0	807	4	09HCB6
11	36	61.0	807	6	09GLX9
12	36	61.0	808	13	042113
13	36	61.0	814	5	09V882
14	36	61.0	831	5	017516
15	36	61.0	991	4	015043
16	36	61.0	1042	4	09H4G6
17	36	61.0	1249	10	023289
18	36	61.0	1631	4	09Y606
19	35	59.3	1637	6	09XSV8

20	35	59.3	4123	4	075851	075851 homo sapien
21	31	52.5	124	2	068137	068137 rhodobacter
22	31	52.5	132	2	056142	056142 salmonella
23	31	52.5	148	4	014583	014583 homo sapien
24	31	52.5	182	4	09H191	09H191 homo sapien
25	31	52.5	195	5	09GZ88	09GZ88 caenorhabdi
26	31	52.5	198	14	09IF37	09IF37 avian infec
27	31	52.5	233	4	099782	099782 homo sapien
28	31	52.5	269	4	09GZ27	09GZ27 homo sapien
29	31	52.5	299	4	09H192	09H192 homo sapien
30	31	52.5	339	14	09DW77	09DW77 rat cytoleg
31	31	52.5	353	10	09SXP5	09SXP5 nicotiana t
32	31	52.5	369	2	046024	046024 corynebacte
33	31	52.5	380	6	09PTV5	09PTV5 bos taurus
34	31	52.5	451	2	006580	006580 mycobacteri
35	31	52.5	463	10	043388	043388 arabidopsis
36	31	52.5	481	14	010651	010651 rubella vir
37	31	52.5	481	14	010652	010652 rubella vir
38	31	52.5	481	14	010653	010653 rubella vir
39	31	52.5	481	14	010654	010654 rubella vir
40	31	52.5	481	14	010655	010655 rubella vir
41	31	52.5	481	14	010656	010656 rubella vir
42	31	52.5	481	14	010657	010657 rubella vir
43	31	52.5	481	14	010658	010658 rubella vir
44	31	52.5	481	14	010659	010659 rubella vir
45	31	52.5	481	14	010660	010660 rubella vir

ALIGNMENTS

RESULT 1
ID 027392 PRELIMINARY; PRT; 204 AA.
AC 027392:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE RANGELIPAIN (FRAGMENT).
OS Trypanosoma rangeli.
OC Eukaryota; Eulendozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5698;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-LDS:
RA Martinez J., Henriksson J., Rydaker M., Gazzulo J.J., Pettersson U.;
RL Mol. Biochem. Parasitol. 129:135-141(1995).
DR EMBL; L38514; AAA79288.1; -;
DR EMBL; L38512; AAA79284.1; -;
DR EMBL; L38513; AAA79286.1; -;
DR HSSP; P25779; 2A1M.
DR MEROPS; C01.075; -;
DR InterPro; IPR000169; -;
DR InterPro; IPR000668; -;
DR Pfam; PF00112; Peptidase_C1; 1.
DR PRINTS; PF00705; PAPA1N.
DR PROSITE; PS00640; THIOI_PROTEASE_ASN; 1.
DR PROSITE; PS00639; THIOI_PROTEASE_HIS; UNKNOWN_1.
KW Hydrolase; Thiol protease.
FT NON_TER 1
SQ SEQUENCE 204 AA; 21602 MM; 006B094AFBF213AD CRC64;

Query Match 62.7%; Score 37; DB 5; Length 204;

Best Local Similarity 83.3%; Pred. No. 12;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 15 CXGPCC 20

Db 104 CSGPC 109

RESULT 2

P71952
ID P71952 PRELIMINARY: PRT: 475 AA.
AC P71952;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE HYPOTHETICAL 51.7 KDA PROTEIN CY441.24C.
GN RV2655C OR MTCY441.24C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxId=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA Skelton J., Churcher C.M., Barrell B.G., Rajandream M.A.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC - SIMILARITY: IN THE C-TERMINUS TO SACCAROPOLYSPORA ERYTHRAEA
CC HYPOTHETICAL 28.9 KDA PROTEIN IN XIS 5 REGION (ORF1).
DR EMBL: Z80225; CAB02331.1; -
DR Tuberculosis; RV2655C; -
KW Hypothetical protein.
SQ SEQUENCE 475 AA; 51718 MW; AC6CF691E5E8FEB CRC64;

Query Match 62.7%; Score 37; DB 2; Length 475;
Best Local Similarity 83.3%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 15 CXGPGC 20
Db 439 CIGPGC 444

RESULT 3
Q9NUM8 PRELIMINARY: PRT: 562 AA.
ID Q9NUM8;
AC Q9NUM8;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE CDNA FLJ11265 FIS, CLONE PLACE1009158.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shitatori A., Sudo H.,
RA Wagatsuna M., Hosolri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Matanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Makatsutsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Niinomiya K., Iwayanagi T.,
RT "NEO human cDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK002127; BAA92094.1; -
DR InterPro: IPR001965; -
DR Pfam: PF00628; PHD; 1.
DR SMART: SM00249; PHD; 1.
SQ SEQUENCE 562 AA; 61425 MW; 3BA79F8DEF9D3D9C CRC64;

Query Match 61.0%; Score 36; DB 4; Length 562;
Best Local Similarity 83.3%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 15 CXGPGC 20
Db 397 CIGPGC 402

RESULT 4
Q9WV00 PRELIMINARY: PRT: 614 AA.
ID Q9WV00;
AC Q9WV00;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE DEATH INDUCER-OBILITERATOR-1 (FRAGMENT).
GN DIO-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99324176; PubMed=10393935;
RA Garcia-Domingo D., Leonardo E., Grandien A., Martinez P., Albar J.P.,
RA Izpisua-Belmonte J.C., Martinez-A.C.;
RT "DIO-1 is a gene involved in onset of apoptosis in vitro, whose
RT misexpression disrupts limb development."
RL Proc. Natl. Acad. Sci. U.S.A. 96:7992-7997(1999).
DR EMBL: AJ238332; CAB48401.1; -
DR InterPro: IPR001965; -
DR Pfam: PF00628; PHD; 1.
DR SMART: SM00249; PHD; 1.
FT NON_TER 614
SQ SEQUENCE 614 AA; 67366 MW; F3E6B897033AADA CRC64;

Query Match 61.0%; Score 36; DB 11; Length 614;
Best Local Similarity 83.3%; Pred. No. 43;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 15 CXGPGC 20
Db 394 CIGPGC 399

RESULT 5
Q94862 PRELIMINARY: PRT: 624 AA.
ID Q94862;
AC Q94862;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE KIAA0762 PROTEIN (FRAGMENT).
GN KIAA0762.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA MEDLINE=99087487; PubMed=9872452;
RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Miyajima N., Tanaka A.,
RA Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XI.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 5:277-286(1998).
DR EMBL: AB018305; BAA34482.1; -
DR InterPro: IPR000884; -
DR Pfam: PF00090; TSP1; 6.
DR SMART: SM00209; TSP1; 1.
FT NON_TER 624
SQ SEQUENCE 624 AA; 70557 MW; 40F2238D29024D03 CRC64;

Query Match 61.0%; Score 36; DB 4; Length 624;
Best Local Similarity 83.3%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 CXGPCC 20
11111
Db 306 CMGPCC 311

RESULT 6
Q9P0U4
ID Q9P0U4 PRELIMINARY: PRT: 656 AA.
AC Q9P0U4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE CPG BINDING PROTEIN.
GN CGBP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20153771; PubMed=10688657;
RA Shin Voo K., Carlone D.L., Jacobsen B.M., Flodin A., Skalniak D.G.;
RT "Cloning of a mammalian transcriptional activator that binds
RT unmethylated CPG motifs and shares a CXXC domain with DNA
RT methyltransferase, human trithorax, and methyl-CPG binding domain
RT protein 1.";
RL Mol. Cell. Biol. 20:2108-2121(2000).
DR EMBL: AF149758; AAF37799.1; -.
DR InterPro: IPR001965; -.
DR Pfam: PF00628; PHD; 1.
DR Pfam: PF02008; zf-CXXC; 1.
DR SMART: SM00249; PHD; 1.
SQ SEQUENCE 656 AA; 75687 MW; AF0C349ECCAD45 CRC64;

Query Match 61.0%; Score 36; DB 4; Length 656;
Best Local Similarity 83.3%; Pred. No. 45;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 CXGPCC 20
11111
Db 375 CMGPCC 380

RESULT 7
Q9P2V7
ID Q9P2V7 PRELIMINARY: PRT: 656 AA.
AC Q9P2V7;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE PROTEIN CONTAINING CXXC DOMAIN 1 (HYPOTHETICAL 75.7 KDA PROTEIN).
GN PCCX1 OR DFEZP434F174.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20261256; PubMed=10799292;
RA Fujino T., Hasegawa M., Shibata S., Kishimoto T., Imai S., Takano T.;
RT "PCCX1, a novel DNA-binding protein with PHD finger and CXXC domain,
RT is regulated by proteolysis.";
RL Biochem. Biophys. Res. Commun. 271:305-310(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Wamutt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB031069; BAA96307.1; -.
DR EMBL: AL136862; CAB66796.1; -.

DR InterPro: IPR001965; -.
DR InterPro: IPR002857; -.
DR Pfam: PF00628; PHD; 1.
DR Pfam: PF02008; zf-CXXC; 1.
DR SMART: SM00249; PHD; 1.
SQ SEQUENCE 656 AA; 75711 MW; 6D2376EA49905A18 CRC64;

Query Match 61.0%; Score 36; DB 4; Length 656;
Best Local Similarity 83.3%; Pred. No. 45;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 CXGPCC 20
11111
Db 375 CMGPCC 380

RESULT 8
Q9W770
ID Q9W770 PRELIMINARY: PRT: 802 AA.
AC Q9W770;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE F-SPONDIN PRECURSOR.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99211385; PubMed=10197528;
RA Debby-Brafman A., Burstyn-Cohen T., Klar A., Kalchauer C.;
RT "F-Spondin, expressed in somite regions avoided by neural crest cells,
RT mediates inhibition of distinct somite domains to neural crest
RT migration.";
RL Neuron 22:475-488(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99325518; PubMed=10399931;
RA Burstyn-Cohen T., Tzarfaty V., Frumkin A., Feinstein Y., Stoeckli E.,
RA Klar A.;
RT "F-spondin is required for accurate pathfinding of commissural axons
RT at the floor plate.";
RL Neuron 23:233-246(1999).
DR EMBL: AF149302; AAD41495.1; -.
DR InterPro: IPR000884; -.
DR InterPro: IPR002861; -.
DR Pfam: PF00090; TSP_1; 6.
DR Pfam: PF02014; Reeler; 1.
DR SMART: SM00209; TSP1; 1.
KW Signal.
FT SIGNAL. 1 20 POTENTIAL.
FT CHAIN 21 802 F-SPONDIN.
SQ SEQUENCE 802 AA; 90509 MW; 0644D2BDD0A0FE12 CRC64;

Query Match 61.0%; Score 36; DB 13; Length 802;
Best Local Similarity 83.3%; Pred. No. 53;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 CXGPCC 20
11111
Db 484 CMGPCC 489

RESULT 9
O42114
ID O42114 PRELIMINARY: PRT: 803 AA.
AC O42114;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)

DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
 DE F-SPONDIN2.
 OS Brachydanio rerio (Zebrafish) (Zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Cypriniformes; Cyprinidae; Rasbora; Danio.
 RN NCB1_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Higashijima S., Nose A., Eguchi G., Hotta Y., Okamoto H.,
 RL Dev. Biol. 0:0-0(1997).
 DR EMBL; AB006087; BAA22811.1; -.
 DR InterPro; IPR000884; -.
 DR Pfam; PF00090; tsp.1; 6.
 DR Pfam; PF02014; Reeler; 1.
 DR SMART; SM00209; TSP1; 1.
 SQ SEQUENCE 803 AA; 90331 MW; 619C8F54B40DC158 CRC64;

Query Match 61.0%; Score 36; DB 13; Length 803;
 Best Local Similarity 83.3%; Pred. No. 53;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 CXGPGC 20
 DB 485 CMGPGC 490

RESULT 10
 O9HCB6 PRELIMINARY; PRT; 807 AA.
 AC O9HCB6:
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DE VSGP/F-SPONDIN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 RN NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Miyamoto K., Morishita Y., Yamazaki M., Minamino N., Kangawa K.,
 RA Matsuo H., Mizutani T., Yamada K., Minegishi T.,
 RT Isolation and characterization of vascular smooth muscle cell growth
 RT promoting factor from bovine ovarian follicular fluid, and its cDNA
 RT cloning from bovine and human ovary."
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB051390; BAB18461.1; -.
 SQ SEQUENCE 807 AA; 90987 MW; 06FC52484206B855 CRC64;

Query Match 61.0%; Score 36; DB 4; Length 807;
 Best Local Similarity 83.3%; Pred. No. 53;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 CXGPGC 20
 DB 489 CMGPGC 494

RESULT 11
 O9GLX9 PRELIMINARY; PRT; 807 AA.
 AC O9GLX9:
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DE VSGP/F-SPONDIN.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.
 OX NCB1_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Miyamoto K., Morishita Y., Yamazaki M., Minamino N., Kangawa K.,
 RA Matsuo H., Mizutani T., Yamada K., Minegishi T.,
 RT Isolation and characterization of vascular smooth muscle cell growth
 RT promoting factor from bovine ovarian follicular fluid, and its cDNA
 RT cloning from bovine and human ovary."
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB051389; BAB18460.1; -.
 SQ SEQUENCE 807 AA; 90976 MW; 4C484B331FB1034C CRC64;

Query Match 61.0%; Score 36; DB 6; Length 807;
 Best Local Similarity 83.3%; Pred. No. 53;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 CXGPGC 20
 DB 489 CMGPGC 494

RESULT 12
 O42113 PRELIMINARY; PRT; 808 AA.
 AC O42113:
 DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
 DE F-SPONDIN.
 OS Brachydanio rerio (Zebrafish) (Zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Cypriniformes; Cyprinidae; Rasbora; Danio.
 RN NCB1_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Higashijima S., Nose A., Eguchi G., Hotta Y., Okamoto H.,
 RL Dev. Biol. 0:0-0(1997).
 DR EMBL; AB006086; BAA22810.1; -.
 DR InterPro; IPR000884; -.
 DR InterPro; IPR002861; -.
 DR Pfam; PF00090; tsp.1; 6.
 DR Pfam; PF02014; Reeler; 1.
 DR SMART; SM00209; TSP1; 1.
 SQ SEQUENCE 808 AA; 90645 MW; 0A24154AA4A89EC7 CRC64;

Query Match 61.0%; Score 36; DB 13; Length 808;
 Best Local Similarity 83.3%; Pred. No. 53;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 CXGPGC 20
 DB 489 CMGPGC 494

RESULT 13
 O9YK82 PRELIMINARY; PRT; 814 AA.
 AC O9YK82:
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
 DE CG6417 PROTEIN.
 CN CG6417.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCB1_TaxID=7227;
 RN [1]

```

RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY:
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaratilake P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abill J.F., Abdayani A., An H.-J., Andrews-Planck C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spletter E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Massaman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhu G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
DR EMBL: AE003636; AAF53197.2; -.
DR FLYBASE: FBgn0032435; CG6417.
SO SEQUENCE 814 AA; 88368 MW; 4CBBDD871ED759FB2 CRC64;

Query Match 61.0%; Score 36; DB 5; Length 814;
Best Local Similarity 83.3%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 CXGPGC 20
DB 35 CCGPGC 40

```

```

RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons N., Percy C., Rifkin L., Roopra A., Saunders D., Showkhen R.,
RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstein L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Brudshaw H., Devlin K.;
RL Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF014939; AAB63927.1; -.
SO SEQUENCE 831 AA; 95398 MW; 1BC6DA0E644264C2 CRC64;

Query Match 61.0%; Score 36; DB 5; Length 831;
Best Local Similarity 83.3%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 CXGPGC 20
DB 192 CCGPGC 197

```

```

RESULT 15
O15043 PRELIMINARY; PRT; 991 AA.
AC O15043:
ID O15043:
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE KIAA0333 (FRAGMENT).
GN KIAA0333.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN:
RX MEDLINE-9734984; PubMed-9205841;
RA Nagase T., Ishikawa K., Nakajima D., Ohira M., Seki N., Miyajima N.,
RA Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 4:141-150(1997).
DR EMBL: AB002331; BAA20791.1; -.
DR InterPro: IPR001052; -.
DR InterPro: IPR001209; -.
DR InterPro: IPR001965; -.
DR Pfam: PF00628; PHD; 1.
DR PRINTS: PR00163; RUBREDOXIN.
DR PROSITE: PS00527; RIBOSOMAL_S14; UNKNOWN_1.
DR SMART: SM00249; PHD; 1.
FT NON TER 1
SO SEQUENCE 991 AA; 107368 MW; 3483D97D3290A7A1 CRC64;

Query Match 61.0%; Score 36; DB 4; Length 991;
Best Local Similarity 83.3%; Pred. No. 63;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 15 CXGPGC 20
| | | |
Db 199 CIGPGC 204

Search completed: November 21, 2001, 15:47:35
Job time: 158 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 15:46:54 ; Search time 22.94 Seconds
(without alignments)
89.656 Million cell updates/sec

Title: US-09-609-721-90

Perfect score: 59
Sequence: 1 XXXXXXXXXXXXXCXCPCXXXXXX 27

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR_68:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	62.7	475	2 C70966	hypothetical prote
2	36	61.0	220	2 A43540	mb-1 protein precu
3	36	61.0	803	2 A47723	F-spondin precurs
4	36	61.0	807	2 A38152	F-spondin - rat
5	36	61.0	831	2 T15108	hypothetical prote
6	36	61.0	1249	2 H71404	hypothetical prote
7	36	61.0	1786	1 MMHGB1	laminin beta-1 cha
8	36	61.0	1786	1 MMMSB1	laminin beta-1 cha
9	35	59.3	373	2 E64310	Iron transport sys
10	31	52.5	124	2 T03574	hypothetical prote
11	31	52.5	132	2 S49160	hypothetical prote
12	31	52.5	195	2 T33188	hypothetical prote
13	31	52.5	389	1 A39429	CAMP response elem
14	31	52.5	451	2 A70539	hypothetical prote
15	31	52.5	486	2 JC4028	activating transcr
16	31	52.5	487	2 G86272	hypothetical prote
17	31	52.5	487	2 F84462	transcription fact
18	31	52.5	505	1 S05380	transcription fact
19	31	52.5	508	2 A45477	CAMP response elem
20	31	52.5	513	2 A24309	genome polypeptid
21	31	52.5	522	1 GNMVRA	structural polypep
22	31	52.5	535	2 S56261	probable membrane
23	31	52.5	641	2 B45871	dnak-type molecu
24	31	52.5	896	2 T27376	hypothetical prote
25	31	52.5	992	1 GNMVRA	structural polypep
26	31	52.5	1063	1 GNMVRA	structural polypep
27	31	52.5	1063	1 GNMVRA	structural polypep
28	31	52.5	1063	1 GNMVRA	structural polypep
29	31	52.5	2946	2 T15840	hypothetical prote

30	30	50.8	96	2 C86649	hypothetical prote
31	30	50.8	133	2 S44579	hypothetical prote
32	30	50.8	155	1 MBMP4	transactivation pr
33	30	50.8	166	2 JW0031	gp delta - satelli
34	30	50.8	172	2 T02055	pathogenesis relat
35	30	50.8	172	2 T04990	pathogenesis-relat
36	30	50.8	217	2 T38385	hypothetical prote
37	30	50.8	235	2 S32747	genome polypeptid
38	30	50.8	247	2 T33654	hypothetical prote
39	30	50.8	282	2 S46722	hypothetical prote
40	30	50.8	291	2 C83504	conserved hypotet
41	30	50.8	300	2 S40837	formate dehydrogen
42	30	50.8	300	2 B86078	formate dehydrogen
43	30	50.8	309	2 C65217	hypothetical 33.8
44	30	50.8	327	2 A46484	apoptosis-mediati
45	30	50.8	441	2 H70632	hypothetical prote

ALIGNMENTS

RESULT 1
C70966
hypothetical protein RV2655c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: C70966
R:Colo, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, R.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, Rajandram, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987
A:Accession: C70966
A:Status: Preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-475 <COL>
A:Cross-references: GB:280225; GB:AL123456; NID:93242265; PIDN:CA802331.1; PID:e26652
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV2655c

Query Match 62.7% Score 37: DB 2; Length 475;
Best Local Similarity 83.3% Pred. No. 28;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 15 CXGPC 20
Db 439 CSGPC 444

RESULT 2
A43540
mb-1 protein precursor - mouse
N:Alternate names: surface Igm complex alpha chain
C:Species: Mus musculus (house mouse)
C:Date: 06-Nov-1992 #sequence_revision 06-Nov-1992 #text_change 05-Nov-1999
C:Accession: A43540; S01648; A60228; A29398; I57521
R:Kashiwamura, S.I.; Koyama, T.; Matsuo, T.; Steinmetz, M.; Kimoto, M.; Sakaguchi, N. J. Immunol. 145, 337-343, 1990
A:Title: Structure of the murine mb-1 gene encoding a putative stigm-associated molecu
A:Reference number: A43540; MUID:90293481
A:Accession: A43540
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-220 <KAS>
A:Cross-references: GB:M31773; NID:q199032; PIDN:AAA39494.1; PID:g387413
R:Sakaguchi, N.; Kashiwamura, S.; Kimoto, M.; Thalmann, P.; Melchers, F. EMBO J. 7, 3457-3464, 1988
A:Title: B lymphocyte lineage-restricted expression of mb-1, a gene with CD3-like str
A:Reference number: S01648; MUID:89091088

A:Accession: S01648
A:Molecule type: mRNA
A:Residues: 1-220 <SAK>
C:Cross-references: EMBL:X13450; NID:g53015; PIDN:CAA31801.1; PID:g53016
R:Hombach, J.; Lottspeich, F.; Reih, M.
Eur. J. Immunol. 20, 2795-2799, 1990
A:Title: Identification of the genes encoding the IgM-alpha and Ig-beta components of the
A:Reference number: A60228; MUID:91099432
A:Accession: A60228
A:Molecule type: protein
A:Residues: 29-38 <HOM>
R:Campbell, K.S.; Hager, E.J.; Friedrich, R.J.; Gambier, J.C.
Proc. Natl. Acad. Sci. U.S.A. 88, 3982-3986, 1991
A:Title: Igm antigen receptor complex contains phosphoprotein products of B29 and mb-1
A:Reference number: A39398; MUID:91219496
A:Accession: A39398
A:Molecule type: protein
A:Residues: 'X', 30-38 <CAM>
R:Travis, A.; Hagman, J.; Grosschedl, R.
Mol. Cell. Biol. 11, 5756-5766, 1991
A:Title: Heterogeneously initiated transcription from the pre-B- and B-cell-specific mb-
A:Reference number: I57521; MUID:92017857
A:Accession: I57521
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-15, 'QA' <RES>
A:Cross-references: GB:S59359; NID:g237080
C:Genetics:
A:Gene: mb-1
C:Keywords: membrane protein
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-220/Product: mb-1 protein #status predicted <MAT>

Query Match 61.0%; Score 36; DB 2; Length 220;
Best Local Similarity 83.3%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 15 CXGPGC 20
Db 21 CGGPGC 26
RESULT 3
A47723
F-spondin precursor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 27-Jun-1994 #sequence_revision 27-Jun-1994 #text_change 17-Nov-2000
C:Accession: A47723
R:Ruiz i Altaba, A.; Cox, C.; Jessell, T.M.; Klar, A.
Proc. Natl. Acad. Sci. U.S.A. 90, 8268-8272, 1993
A:Title: Ectopic neural expression of a floor plate marker in frog embryos injected with
A:Reference number: A47723; MUID:9336785
A:Accession: A47723
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1803 <RUI>
A:Cross-references: GB:U09123; NID:g409244; PIDN:AAA19105.1; PID:g409245
C:Superfamily: F-spondin; thrombospondin type 1 repeat homology
F:435-489/Domain: thrombospondin type 1 repeat homology <THR2>
F:607-662/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 61.0%; Score 36; DB 2; Length 803;
Best Local Similarity 83.3%; Pred. No. 62;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 15 CXGPGC 20
Db 483 CXGPGC 488
RESULT 4

A38152
F-spondin - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 17-Nov-2000
C:Accession: A38152
R:Klar, A.; Balassare, M.; Jessell, T.M.
Cell 69, 95-110, 1992
A:Title: F-spondin: a gene expressed at high levels in the floor plate encodes a secr
A:Reference number: A38152; MUID:92208952
A:Accession: A38152
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-807 <KLA>
A:Cross-references: GB:M88469; NID:g204176; PIDN:AAA41174.1; PID:g204177
A:Experimental source: embryo floor plate
A:Note: sequence extracted from NCBI backbone (NCBIN:90877, NCBI:P:90878)
C:Superfamily: F-spondin; thrombospondin type 1 repeat homology
F:441-495/Domain: thrombospondin type 1 repeat homology <THR2>
F:500-555/Domain: thrombospondin type 1 repeat homology <THR3>
F:557-611/Domain: thrombospondin type 1 repeat homology <THR4>
F:613-666/Domain: thrombospondin type 1 repeat homology <THR1>
F:667-721/Domain: thrombospondin type 1 repeat homology <THR5>
F:753-807/Domain: thrombospondin type 1 repeat homology <THR6>

Query Match 61.0%; Score 36; DB 2; Length 807;
Best Local Similarity 83.3%; Pred. No. 62;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 15 CXGPGC 20
Db 489 CXGPGC 494

RESULT 5
T15108
Hypothetical protein ZC132.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15108
R:Bradshaw, H.; Devlin, K.
submitted to the EMBL Data Library, July 1997
A:Description: The sequence of C. elegans cosmid ZC132.
A:Reference number: Z18294
A:Accession: T15108
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-831 <BRA>
A:Cross-references: EMBL:AF014939; NID:g2275620; PID:g2275624; PIDN:AAB63927.1; GSPDB
A:Experimental source: strain Bristol N2; clone ZC132
C:Genetics:
A:Gene: CESP:ZC132.4
A:Map position: 5
A:Introns: 432/3; 543/3

Query Match 61.0%; Score 36; DB 2; Length 831;
Best Local Similarity 83.3%; Pred. No. 64;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 15 CXGPGC 20
Db 192 CGGPGC 197

RESULT 6
H71404
Hypothetical protein d13195c - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
A:Variety: Columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 20-Jun-2000
C:Accession: H71404
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; D

P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terryn, N.; Giel
 arangh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Scheffler, M.; Funk, B.
 Nature 331, 485-488, 1998
 A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
 erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Reckman, S.; Aus
 C.; Chalvatzis, N.
 A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
 A:Reference number: A11400; MUID:98121113
 A:Accession: H71404
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1249 <BEV>
 A:Cross-references: GB:297335; NID:92244747; PID:92244787
 C:Genetics:
 A:Map position: 4COP9-4G3845
 C:Superfamily: Arabidopsis thaliana hypothetical protein d13195c

Query Match	Best Local Similarity	Score 36;	DB 2;	Length 1249;
Matches 5; Conservative	83.3%;	Pred. No. 87;	1;	Indels 0; Gaps 0;
Mismatches	0;	Mismatches	0;	

QY 15 CXGPCC 20
 1 |||||
 Db 12 CGGPCC 17

RESULT 7
 MMHUB1
 laminin beta-1 chain precursor - human
 N:Alternate names: laminin chain B1
 C:Species: Homo sapiens (man)
 C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 19-Jan-2001
 C:Accession: S13547; A28483; A26994; S23566
 R:Vuolteenaho, R.; Chow, L.T.; Tryggvason, K.
 J. Biol. Chem. 265, 15611-15616, 1990
 A:Title: Structure of the human laminin B1 chain gene.
 A:Reference number: S13547; MUID:90368768
 A:Accession: S13547
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1786 <VUD>
 A:Cross-references: GB:M61951; GB:J02778; NID:9186911; PIDN:AAAS9486.1; PID:9186913
 A:Note: The nucleotide sequence was submitted to Genbank, February 1991
 R:Pikkarainen, T.; Eddy, R.; Fukushima, Y.; Byers, M.; Shows, T.; Pihlajaniemi, T.; Sare
 J. Biol. Chem. 262, 10454-10462, 1987
 A:Title: Human laminin B1 chain. A multidomain protein with gene (LAMB1) locus in the q2
 A:Reference number: A28483; MUID:87280097
 A:Accession: A28483
 A:Molecule type: mRNA
 A:Residues: 1-1786 <PIK>
 A:Cross-references: GB:M61951; GB:J02778; NID:9186911; PIDN:AAAS9486.1; PID:9186913
 R:Jaye, M.; Modi, W.S.; Ricca, G.A.; Mudd, R.; Chiu, I.M.; O'Brien, S.J.; Drohan, W.N.
 Am. J. Hum. Genet. 41, 605-615, 1987
 A:Title: Isolation of a cDNA clone for the human laminin-B1 chain and its gene localizat
 A:Reference number: A26994; MUID:88021029
 A:Accession: A26994
 A:Molecule type: mRNA
 A:Residues: 1276-1469, 'V', 1471-1695, 'G', 1697-1709 <JAV>
 A:Cross-references: EMBL:M20206; NID:9186914; PIDN:AAAS9487.1; PID:9186915
 R:Vuolteenaho, R.; Kallunki, T.; Chow, L.; Ikonen, T.; Pihlajaniemi, K.
 In Extracellular Matrix Genes, Sandell L.J. and Boyd C.D., eds., pp. 175-193, Academic P
 A:Title: Genes for the human laminin B1 and B2 chains.
 A:Reference number: S23566
 A:Accession: S23566
 A:Molecule type: DNA
 A:Residues: 762-1786 <VUD>
 A:Note: mRNA was also sequenced
 C:Genetics:
 A:Gene: GDB:LAMB1
 A:Cross-references: GDB:119357; OMIM:150240
 A:Map position: 7q31.1-7q31.3
 A:Introns: 13/1; 71/3; 117/1; 141/3; 204/3; 226/1; 293/3; 334/1; 397/1; 457/1; 494/3; 52

64/3; 1513/1; 1582/2; 1629/3; 1688/3; 1742/1
 C:Complex: Laminins are trimers of an alpha-type, a beta-type, and a gamma-type lamin
 C:Function:
 A:Description: Interact with cells and with other basement membrane proteins to promo
 C:Superfamily: laminin beta-1 chain; laminin-type EGF-like homology
 C:Keywords: basement membrane; calcium binding; cell binding; coiled coil; extracellu
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-1786/Product: laminin beta-1 chain #status predicted <MAT>
 F:22-270/Domain: VI <DOM6>
 F:221-548/Domain: V <DOM5>
 F:271-332/Domain: laminin-type EGF-like homology <LE01>
 F:335-395/Domain: laminin-type EGF-like homology <LE02>
 F:398-455/Domain: laminin-type EGF-like homology <LE03>
 F:458-507/Domain: laminin-type EGF-like homology <LE04>
 F:463-468/Region: cell adhesion #status predicted
 F:510-540/Domain: laminin-type EGF-like homology #status atypical <LE05>
 F:549-774/Domain: IV <DOM4>
 F:662-668/Region: cell adhesion #status predicted
 F:772-818/Domain: laminin-type EGF-like homology <LE06>
 F:775-1178/Domain: III <DOM3>
 F:821-864/Domain: laminin-type EGF-like homology <LE07>
 F:867-914/Domain: laminin-type EGF-like homology <LE08>
 F:917-973/Domain: laminin-type EGF-like homology <LE09>
 F:923-927/Region: cell adhesion #status predicted
 F:950-954/Region: cell adhesion #status predicted
 F:976-1025/Domain: laminin-type EGF-like homology <LE10>
 F:1028-1081/Domain: laminin-type EGF-like homology <LE11>
 F:1084-1129/Domain: laminin-type EGF-like homology <LE12>
 F:1132-1176/Domain: laminin-type EGF-like homology <LE13>
 F:1179-1397/Domain: II <DOM2>
 F:1179-1397/Region: heptad repeats
 F:1398-1430/Domain: alpha <ALP>
 F:1431-1786/Domain: I <DOM1>
 F:1431-1786/Region: heptad repeats
 F:30-35/Disulfide bonds: #status predicted
 F:120,356,519,577,965,1041,1195,1279,1336,1343,1487,1542,1643/Binding site: carboxydr
 F:1179,1182,1785/Disulfide bonds: interchain #status predicted

Query Match	Best Local Similarity	Score 36;	DB 1;	Length 1786;
Matches 5; Conservative	83.3%;	Pred. No. 1,1e+02;	1;	Indels 0; Gaps 0;
Mismatches	0;	Mismatches	1;	

QY 15 CXGPCC 20
 1 |||||
 Db 1425 CGGPCC 1430

RESULT 8
 MMMSB1
 laminin beta-1 chain precursor - mouse
 N:Alternate names: laminin chain B1
 C:Species: Mus musculus (house mouse)
 C:Date: 28-Feb-1996 #sequence_revision 30-Jun-1991 #text_change 10-Dec-1999
 C:Accession: A26413; S02679; S05326; S14877; A02871; S02036; S13543
 R:Sasaki, M.; Kato, S.; Kohno, K.; Martin, G.R.; Yamada, Y.
 Proc. Natl. Acad. Sci. U.S.A. 84, 935-939, 1987
 A:Title: Sequence of the cDNA encoding the laminin B1 chain reveals a multidomain pro
 A:Reference number: A26413; MUID:87147212
 A:Accession: A26413
 A:Molecule type: mRNA
 A:Residues: 1-1786 <SNS>
 A:Cross-references: EMBL:M15525; NID:9198700
 A:Note: translation in Genbank has additional 48 residues at the amino end
 R:Fujimura, S.; Shinkai, H.; Deutzmann, R.; Paulsson, M.; Timpl, R.
 Biochem. J. 252, 453-461, 1988
 A:Title: Structure and distribution of N-linked oligosaccharide chains on various dom
 A:Reference number: S02678; MUID:88326259
 A:Accession: S02679
 A:Molecule type: protein
 A:Residues: 28-42;932-946 <FUJ>
 R:Hartl, L.; Oberhauser, I.; Deutzmann, R.
 Eur. J. Biochem. 173, 629-635, 1988

A:Title: The N terminus of laminin A chain is homologous to the B chains.
A:Reference number: S00624; MUID:88225080
A:Accession: S05326
A:Molecule type: protein
A:Residues: 457-466;854-868;932-946 <HAR>
R:Mann, K.; Deutzmann, R.; Timpl, R.
Eur. J. Biochem. 178, 71-80, 1988
A:Title: Characterization of proteolytic fragments of the laminin-5/6 complex and th
A:Reference number: S08895; MUID:89078415
A:Accession: S14877
A:Molecule type: protein
A:Residues: 350-620 <MAN>
R:Barlow, D.P.; Green, N.M.; Kurkinen, M.; Hogan, B.L.M.
EMBO J. 3, 2355-2362, 1984
A:Title: Sequencing of laminin B chain cDNAs reveals C-terminal regions of coiled-coil a
A:Reference number: A02870; MUID:85051302
A:Accession: A02871
A:Molecule type: mRNA
A:Residues: 1292-1530, MEMP, 1535-1691, C', 1693-1748, N', 1750-1786 <BAR>
A:Cross-references: EMBL:X05212; NID:952861; PIDN:CAA28833.1; PID:9809042
R:Deutzmann, R.; Huber, J.; Schmetz, K.A.; Oberbauer, I.; Hartl, L.
Eur. J. Biochem. 177, 35-45, 1988
A:Title: Structural study of long arm fragments of laminin. Evidence for repetitive C-te
A:Reference number: S01790; MUID:89030693
A:Accession: S02036
A:Molecule type: protein
A:Residues: 1561-1587 <DEU>
R:Paulsson, M.; Deutzmann, R.; Timpl, R.; Dalzoppo, D.; Oermatt, E.; Engel, J.
EMBO J. 4, 309-316, 1985
A:Title: Evidence for coiled-coil alpha-helical regions in the long arm of laminin.
A:Reference number: S13543; MUID:85257455
A:Accession: S13543
A:Molecule type: protein
A:Residues: 1700-1748, N', 1750-1759 <PAU>
C:Genetics:
A:Gene: Lamb-1
A:Map position: 12
C:Complex: Laminins are trimers of an alpha-type, a beta-type, and a gamma-type laminin
C:Function:
A:Description: Interact with cells and with other basement membrane proteins to promote
C:Superfamily: Laminin beta-1 chain; Laminin-type EGF-like homology
C:Keywords: basement membrane; calcium binding; cell binding; coiled coil; extracellular
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-1786/Product: laminin beta-1 chain #status predicted <MAN>
F:22-270/Domain: VI <DOM6>
F:271-540/Domain: V <DOM5>
F:271-532/Domain: laminin-type EGF-like homology <LE01>
F:335-395/Domain: laminin-type EGF-like homology <LE02>
F:398-455/Domain: laminin-type EGF-like homology <LE03>
F:458-507/Domain: laminin-type EGF-like homology <LE04>
F:510-540/Domain: laminin-type EGF-like homology #status atypical <LE05>
F:541-772/Domain: IV <DOM4>
F:773-1184/Domain: III <DOM3>
F:773-818/Domain: laminin-type EGF-like homology <LE06>
F:821-864/Domain: laminin-type EGF-like homology <LE07>
F:867-914/Domain: laminin-type EGF-like homology <LE08>
F:917-973/Domain: laminin-type EGF-like homology <LE09>
F:976-1026/Domain: laminin-type EGF-like homology <LE10>
F:1028-1061/Domain: laminin-type EGF-like homology <LE11>
F:1084-1129/Domain: laminin-type EGF-like homology <LE12>
F:1132-1176/Domain: laminin-type EGF-like homology <LE13>
F:1183-1397/Domain: II <DOM2>
F:1183-1397/Region: heptad repeats
F:1398-1430/Domain: alpha <ALP>
F:1431-1786/Region: heptad repeats
F:1431-1786/Domain: I <DOM1>
F:22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F:30-35/Disulfide bonds: #status predicted
F:120, 356, 519, 677, 1041, 1195, 1279, 1336, 1343, 1487, 1533, 1542, 1643/Binding site: carbohydrat
F:1179, 1182, 1785/Disulfide bonds: Interchain #status predicted

Query Match

61.0%; Score 36; DB 1; Length 1786;

Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 15 CXGPGC 20
Db 1425 CGGPGC 1430
RESULT 9
E64310
iron transport system binding protein - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: E64310
R:Bull, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak
; Reich, C.I.; Overbeek, R.; Kirsness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodet,
; rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc
A:Reference number: A64300; MUID:96337999
A:Accession: E64310
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-373 <BUL>
A:Cross-references: GB:U67466; GB:L77117; NID:g1590867; PIDN:AAB98067.1; PID:g1590868
C:Genetics:
A:Map position: FOR79630-80751
Query Match 59.3%; Score 35; DB 2; Length 373;
Best Local Similarity 83.3%; Pred. No. 51;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 15 CXGPGC 20
Db 61 CCGPGC 66
RESULT 10
T03574
hypothetical protein - Rhodobacter capsulatus
C:Species: Rhodobacter capsulatus
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000
C:Accession: T03574
R:Vleck, C.; Paces, V.; Maltsev, N.; Paces, J.; Haselkorn, R.; Fomstein, M.
Proc. Natl. Acad. Sci. U.S.A. 94, 9384-9388, 1997
A:Title: Sequence of a 189-kb segment of the chromosome of Rhodobacter capsulatus SB1
A:Reference number: Z14955; MUID:97404404
A:Accession: T03574
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-124 <VLC>
A:Cross-references: EMBL:AF010496; NID:g3128256; PIDN:AAC16227.1; PID:g3128375
C:Genetics:
A:Map position: 1
C:Superfamily: Bacillus phage phi-105 hypothetical protein 19
Query Match 52.5%; Score 31; DB 2; Length 124;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 15 CXGPGC 20
Db 7 CSAPGC 12
RESULT 11
S49160
hypothetical protein 14.5K - Salmonella typhimurium
C:Species: Salmonella typhimurium
C:Date: 16-Feb-1995 #sequence_revision 14-Jul-1995 #text_change 24-Nov-1999

C:Accession: S49160
R:Nieweg, A.; Bremer, E.
submitted to the EMBL Data Library, October 1993
A:Description: Molecular characterization of the structural genes (tsx) for the nucleoside
A:Reference number: S49142
A:Accession: S49160
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-132 <NIE>
A:Cross-references: EMBL:22657; NID:9510211; PIDN:CAA81399.1; PID:9510213
C:Superfamily: Salmonella typhimurium hypothetical protein 14.5k

Query Match 52.5%; Score 31; DB 2; Length 132;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 15 CXGPC 20
DB 98 CASPGC 103

RESULT 12
T33188
hypothetical protein F22F7.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 20-Jun-2000
C:Accession: T33188
R:Miller, N.; Kramer, J.; Smith, A.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of C. elegans cosmid F22F7.
A:Reference number: Z21299
A:Accession: T33188
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-195 <MIL>
A:Cross-references: EMBL:AF067937; PIDN:AC19220.1; GSPDB:GN00023; CESP:F22F7.7
A:Experimental source: strain Bristol N2; clone F22F7
C:Genetics:
A:Gene: CESP:F22F7.7
A:Map position: 5
A:introns: 57/3; 96/3; 160/2
C:Superfamily: Schizosaccharomyces pombe hypothetical protein SPBC31F10.03

Query Match 52.5%; Score 31; DB 2; Length 195;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 15 CXGPC 20
DB 154 CAGPNC 159

RESULT 13
A39429
cAMP response element-binding protein ATF2 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A39429
R:Kageyama, R.; Sasai, Y.; Nakamishi, S.
J. Biol. Chem. 266, 15525-15531, 1991
A:Title: Molecular characterization of transcription factors that bind to the cAMP response
A:Reference number: A39429; MUID:91332085
A:Accession: A39429
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-389 <KAG>
A:Cross-references: GB:M65148; NID:9206569; PIDN:AAA2013.1; PID:9206570
C:Superfamily: cAMP response element-binding protein 1; fos/jun DNA-binding domain homol
C:Keywords: DNA binding; nucleus; transcription regulation
F:231-271/Domain: fos/jun DNA-binding domain homology <FJD>

Query Match 52.5%; Score 31; DB 1; Length 389;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 15 CXGPC 20
DB 9 CTAPGC 14

RESULT 14
A70539
hypothetical protein RV1128c - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: A70539
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987
A:Accession: A70539
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-451 <COL>
A:Cross-references: GB:295585; GB:AL22456; NID:93261787; PIDN:CAB09039.1; PID:921172
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: RV1128c
C:Superfamily: Mycobacterium tuberculosis hypothetical protein RV1945

Query Match 52.5%; Score 31; DB 2; Length 451;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 15 CXGPC 20
DB 363 CSAPGC 368

RESULT 15
JC4028
activating transcription factor 2 - African clawed frog
N:Alternate names: cyclic AMP-response element-binding protein
C:Species: Xenopus laevis (African clawed frog)
C:Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 26-Aug-1999
C:Accession: JC4028
R:Villarreal, X.C.; Richter, J.D.
Gene 153, 225-229, 1995
A:Title: Analysis of ATF2 gene expression during early Xenopus laevis developme nt.
A:Reference number: JC4028; MUID:95180723
A:Accession: JC4028
A:Molecule type: mRNA
A:Residues: 1-486 <VIL>
A:Cross-references: GB:U16158; NID:9887779; PIDN:AAA69518.1; PID:9710326
C:Comment: This protein is a sequence-specific DNA-binding protein that mediates tran
C:Superfamily: cAMP response element-binding protein 1; fos/jun DNA-binding domain ho
C:Keywords: leucine zipper; phosphoprotein; transcription regulation
F:337-398/Region: leucine zipper motif
F:332-359/Domain: fos/jun DNA-binding domain homology <FJD>
F:332-359/Region: basic
F:332-359/Region: basic
F:332-359/Region: basic
F:102/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
F:102/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match 52.5%; Score 31; DB 2; Length 486;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 15 CXGPC 20

Db 9 CTAPGC 14

Search completed: November 21, 2001, 15:46:55
Job time: 118 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 15:46:24 ; Search time 12.78 Seconds
(without alignments)
72.371 Million cell updates/sec

Title: US-09-609-721-90

Perfect score: 59

Sequence: 1 XXXXXXXXXXXXXCXPCCXXXXXX 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	36	61.0	220 1 C79A_MOUSE	P11911 mus musculus
2	36	61.0	803 1 FSP0_XENLA	P35447 xenopus lae
3	36	61.0	807 1 FSP0_RAT	P35446 ratius norv
4	36	61.0	867 1 SSP0_BOVIN	P98167 bos taurus
5	36	61.0	1786 1 LMB1_HUMAN	P07942 homo sapien
6	36	61.0	1786 1 LMB1_MOUSE	P24659 mus musculus
7	35	59.3	373 1 Y085_MERJA	Q57550 methanococ
8	31	52.5	487 1 ATF2_CHICK	Q93602 gallus gall
9	31	52.5	487 1 ATF2_HUMAN	P15336 homo sapien
10	31	52.5	487 1 ATF2_MOUSE	P16951 mus musculus
11	31	52.5	487 1 ATF2_RAT	P00969 ratius norv
12	31	52.5	522 1 POLS_RUBV	P08564 rubella vir
13	31	52.5	535 1 YFHE_YEAST	P43590 saccharomyc
14	31	52.5	641 1 H57H_HUMAN	P34931 homo sapien
15	31	52.5	992 1 POLS_RUBVM	P08563 rubella vir
16	31	52.5	1063 1 POLS_RUBVM	P21480 rubella vir
17	31	52.5	1063 1 POLS_RUBVM	P19755 rubella vir
18	31	52.5	1233 1 NME3_HUMAN	P07566 rubella vir
19	31	52.5	1233 1 NME3_HUMAN	Q14957 homo sapien
20	31	52.5	1435 1 Y194_HUMAN	Q12766 homo sapien
21	31	50.8	166 1 VDEL_BPP4	P12551 bacterioph
22	30	50.8	282 1 YHPO_YEAST	P38806 saccharomyc
23	30	50.8	300 1 FDOH_ECOLI	P32175 escherichia
24	30	50.8	309 1 ALSK_ECOLI	P32718 escherichia
25	30	50.8	324 1 FASA_MOUSE	P25446 mus musculus
26	30	50.8	397 1 ATF7_HUMAN	P17544 homo sapien
27	30	50.8	653 1 RAE2_HUMAN	P24366 homo sapien
28	30	50.8	656 1 RAE2_HUMAN	P26314 homo sapien
29	30	50.8	775 1 TNP3_MOUSE	Q60769 mus musculus
30	30	50.8	790 1 TNP3_HUMAN	P21580 homo sapien
31	30	50.8	928 1 YAO8_HUMAN	Q9Y211 homo sapien
32	30	50.8	1377 1 CID_DROME	P19538 drosophila
33	30	50.8	1441 1 VGLM_BUNLV	P09612 bunyavirus

34	30	50.8	1441 1 VGLM_BUNSH	P04875 bunyavirus
35	30	50.8	2731 1 RRPB_CWMTH	P29982 murine coro
36	30	50.8	2733 1 RRPB_CWMAS	P16342 murine coro
37	29	49.2	56 1 Y6KD_STRLI	P22400 streptomyce
38	29	49.2	72 1 MT11_MYTED	P80246 mytilus edu
39	29	49.2	72 1 MT12_MYTED	P80247 mytilus edu
40	29	49.2	72 1 MT13_MYTED	P80248 mytilus edu
41	29	49.2	72 1 MT14_MYTED	P80249 mytilus edu
42	29	49.2	94 1 RL43_SCHPO	O94686 schistosom
43	29	49.2	207 1 EGG2_SCHUA	P19469 schistosoma
44	29	49.2	212 1 EGG1_SCHUA	P19470 schistosoma
45	29	49.2	294 1 FDNH_ECOLI	P24184 escherichia

ALIGNMENTS

RESULT 1
C79A_MOUSE STANDARD: PRT: 220 AA.
AC P11911;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE B-CELL ANTIGEN RECEPTOR COMPLEX ASSOCIATED PROTEIN ALPHA-CHAIN
DE PRECURSOR (IG-ALPHA) (MB-1 MEMBRANE GLYCOPROTEIN) (SURFACE-IGM-
DE ASSOCIATED PROTEIN) (MEMBRANE-BOUND IMMUNOGLOBULIN ASSOCIATED
DE PROTEIN) (CD79A).
CN CD79A OR IGA OR MB-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X DBA/2J;
RX MEDLINE=89091088; PubMed=2463161;
RA Sakaguchi N., Kashiwamura S., Kimoto M., Thalmann P., Melchers F.;
RT "B lymphocyte lineage-restricted expression of mb-1, a gene with CD3-
RT like structural properties.";
RL EMBO J. 7:3457-3464(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BA1B/C; TISSUE=Liver;
RX MEDLINE=90293481; PubMed=2358676;
RA Kashiwamura S.-I., Koyama T., Matsuo T., Steinmetz M., Kimoto M.,
RT Sakaguchi N.;
RT "Structure of the murine mb-1 gene encoding a putative
RT sigw-associated molecule.";
RL J. Immunol. 145:337-343(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC MEDLINE=92347937; PubMed=1639443;
RA Flaiswinkel H., Reith M.;
RT "Antigenic cloning of the Ig-alpha subunit of the human B-cell
RT molecular receptor complex.";
RL Immunogenetics 36:266-269(1992).
CC - FUNCTION: ASSOCIATED TO SURFACE IGM-RECEPTOR; MAY BE INVOLVED IN
CC SIGNAL TRANSDUCTION.
CC - SUBUNIT: HETERODIMER OF ALPHA AND BETA CHAINS, DISULFIDE-LINKED.
CC - SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC - TISSUE SPECIFICITY: B-CELLS.
CC - PTM: PHOSPHORYLATED ON TYROSINE AS A RESULT OF B-CELL ACTIVATION.
CC - SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR	EMBL: X13450; CAA31801.1; -	DR	EMBL: M31773; AAA39494.1; -
DR	PIR: S01648; S01648.	DR	PIR: A43540; A43540.
DR	PIR: A43540; A43540.	DR	MEG: MGI:101774; Cd79a.
KW	Immunoglobulin domain; Transmembrane; Glycoprotein; B-cell; Signal; Phosphorylation.	KW	Immunoglobulin domain; Transmembrane; Glycoprotein; B-cell; Signal; Phosphorylation.
FT	SIGNAL	FT	SIGNAL
FT	CHAIN	FT	CHAIN
FT	DOMAIN	FT	DOMAIN
FT	TRANSMEM	FT	TRANSMEM
FT	DOMAIN	FT	DOMAIN
FT	DISULFID	FT	DISULFID
FT	DISULFID	FT	DISULFID
FT	CARBOHYD	FT	CARBOHYD
FT	CONFLICT	FT	CONFLICT
SO	SEQUENCE	SO	SEQUENCE
	220 AA; 24261 MW; A654123C8177B29 CRO64;		220 AA; 24261 MW; A654123C8177B29 CRO64;

Query Match	61.0%;	Score 36;	DB 1;	Length 220;
Best Local Similarity	83.3%;	Pred. No. 8;		
Matches	5;	Conservative	0;	Mismatches 1;
				Indels 0;
				Gaps 0;

QY	15	CXGPGC	20
Db	21	CLGPGC	26

RESULT	2
FSPO_XENLA	
ID_FSPO_XENLA	STANDARD;
	PRT; 803 AA

DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE F-SPONDIN PRECURSOR.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_Taxid=8355;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=93376785; PubMed=8367492;
 RA Ruiz I Altaba A., Cox C., Jessell T.M., Klar A.;
 RT "Ectopic neural expression of a floor plate marker in frog embryos
 RL injected with the midline transcription factor Pintlavlav1".
 RL Proc. Natl. Acad. Sci. U.S.A. 90:8266-8272(1993).
 CC -I- FUNCTION: PROMOTES THE ATTACHMENT OF SPINAL CORD AND SENSORY
 CC NEURON CELLS AND THE OUTGROWTH OF NUTRIENTS IN VITRO. MAY
 CC CONTRIBUTE TO THE GROWTH AND GUIDANCE OF AXONS IN BOTH THE SPINAL
 CC CORD AND THE PNS.
 CC -I- SUBCELLULAR LOCATION: SECRETED.
 CC -I- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN THE FLOOR PLATE.
 CC -I- SIMILARITY: CONTAINS 6 TYPE-1 TSP DOMAINS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L09123; AAA19105.1; -;
 DR PIR: A47723; A47723;
 DR InterPro: IPR0000884; -;
 DR InterPro: IPR002861; -;
 DR Pfam: PF02014; Reeler; 1.

DR	Pfam: PF00090; tsp.1; 6.
DR	PROSITE; PSS0092; TSP1; 6.
KW	Glycoprotein; Signal; Repeat; Cell adhesion.
FT	SIGNAL 1 23 POTENTIAL.
FT	CHAIN 24 803 F-SPONDIN.
FT	DOMAIN 437 488 TSP TYPE-1 1.
FT	DOMAIN 496 548 TSP TYPE-1 2.
FT	DOMAIN 553 604 TSP TYPE-1 3.
FT	DOMAIN 609 661 TSP TYPE-1 4.
FT	DOMAIN 665 716 TSP TYPE-1 5.
FT	DOMAIN 751 803 TSP TYPE-1 6.
FT	CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD 677 677 N-LINKED (GLCNAC. . .) (POTENTIAL)
SO	SEQUENCE 803 AA; 90702 MM; D3A4E329548AD9 C6R64;

Query Match	61.0%:	Score 36:	DB 1:	Length 803:
Best Local Similarity	83.3%:	Pred. No. 21:		
Matches	5:	Conservative	0:	Mismatches 1:
				Indels 0:
				Gaps 0:

QY	15	CXGPGC	20
Db	483	CMGPGC	488

RESULT	3
FSPO_RAT	
ID	FSPO_RAT
STANDARD;	
PRT;	807 AA

DT	01-JUN-1994	(Rel. 29, Created)
DT	01-JUN-1994	(Rel. 29, Last sequence update)
DT	01-OCT-2000	(Rel. 40, Last annotation update)
DE	F-SPONDIN	PRECURSOR.
OS	Rattus norvegicus	(Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
OX	NCBI_Taxid=10116;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Embryonic floor plate;	
RX	MEDLINE=92208952; PubMed=1555244;	
RA	Klar A., Balassare M., Jessell T.M.;	
RT	F."spodin": a gene expressed at high levels in the floor plate	
RT	encodes a secreted protein that promotes neural cell adhesion and	
RT	neurite extension."	
RL	Cell 69:95-110(1992).	
CC	-1- FUNCTION: PROMOTES THE ATTACHMENT OF SPINAL CORD AND SENSOR	
CC	NEURON CELLS AND THE OUTGROWTH OF NEURITES IN VITRO. MAY	
CC	CONTRIBUTE TO THE GROWTH AND GUIDANCE OF AXONS IN BOTH THE SPINAL	
CC	CORD AND THE PNS.	
CC	-1- SUBCELLULAR LOCATION: SECRETED.	
CC	-1- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN THE FLOOR PLATE.	
CC	-1- SIMILARITY: CONTAINS 6 TYPE-1 TSP DOMAINS.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL, M88469, AAA41174.1; -.	
DR	PIR: A38152, A38152.	
DR	InterPro: IPR000884; -.	
DR	InterPro: IPR002861; -.	
DR	Pfam: PF02014; Reeler; 1.	
DR	Pfam: PF00090; tsp_1; 6.	
DR	PROSITE: PS50092; TSP1; 5.	
KM	Glycoprotein; Signal; Repeat; Cell adhesion.	
FT	SIGNAL	1 28
FT	CHAIN	29 807
FT	DOMAIN	443 494
FT		F-SPONDIN.
FT		TSP TYPE-1 1.

```

FT DOMAIN 502 554 TSP TYPE-1 2.
FT DOMAIN 559 610 TSP TYPE-1 3.
FT DOMAIN 615 665 TSP TYPE-1 4.
FT DOMAIN 669 720 TSP TYPE-1 5.
FT DOMAIN 755 807 TSP TYPE-1 6.
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 681 681 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 807 AA; 90773 MW; 309525F9EAFE89A CRC64;

Query Match
Best Local Similarity 83.3%; Score 36; DB 1; Length 807;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 15 CXGPC 20
Db 489 CMGPC 494

RESULT 4
SSPO_BOVIN STANDARD; PRT; 867 AA.
AC P98167;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE SCO-SPONDIN (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Endomycyte;
RX MEDLINE=96338614; PubMed=8743952;
RA Gobron S., Monnerie H., Meinzel R., Creveaux I., Lehmann W.,
RA Lamalle D., Dastugue B., Meinzel A.;
RA "SCO-spondin: a new member of the thrombospondin family secreted by
RT the subcommissural organ is a candidate in the modulation of neuronal
RT aggregation."
RT J. Cell Sci. 109:1053-1061(1996).
RL
CC -I- FUNCTION: INVOLVED IN THE MODULATION OF NEURONAL AGGREGATION.
CC -I- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -I- TISSUE SPECIFICITY: SUBCOMMISSURAL ORGAN.
CC -I- DEVELOPMENTAL STAGE: EMBRYO.
CC -I- SIMILARITY: BELONGS TO THE THROMBOSPONDIN FAMILY.
CC -I- SIMILARITY: CONTAINS AT LEAST 4 TSP TYPE-1 DOMAINS.
CC -I- SIMILARITY: CONTAINS AT LEAST 2 EGF-LIKE DOMAINS.
CC -I- SIMILARITY: CONTAINS AT LEAST 1 F5/8 TYPE C DOMAIN.
CC -I- SIMILARITY: CONTAINS AT LEAST 3 LDL-RECEPTOR CLASS A DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC
CC EMBL: X93922; CAA63815.1;
CC DR HSSP: P01130; IAJT.
CC DR InterPro: IPR000421;
CC DR InterPro: IPR000884;
CC DR InterPro: IPR00107;
CC DR InterPro: IPR002172;
CC DR InterPro: IPR002919;
CC DR Pfam: PF00754; F5_F8_Type_C; 1.
CC DR Pfam: PF01826; TIL; 1.
CC DR Pfam: PF00057; Idl_recept_a; 3.
CC DR Pfam: PF00090; tsp_1; 4.
CC DR Pfam: PF00093; vwc; 1.
CC DR PROSITE: PS01285; FA58C_1; 1.

```

```

DR PROSITE: PS01286; FA58C_2; 1.
DR PROSITE: PS01209; LDLRA_1; 3.
DR PROSITE: PS50068; LDLRA_2; 3.
DR PROSITE: PS50092; TSP1; 4.
KW Glycoprotein; Cell adhesion; Calcium-binding; Repeat; EGF-like domain.
FT DOMAIN 1 1 TSP TYPE-1 1.
FT DOMAIN 26 81 EGF-LIKE 1.
FT DOMAIN 103 142 EGF-LIKE 2.
FT DOMAIN 143 180 TSP TYPE-1 2.
FT DOMAIN 185 243 F5/8 TYPE C.
FT DOMAIN 344 502 LDL-RECEPTOR CLASS A 1.
FT DOMAIN 506 544 LDL-RECEPTOR CLASS A 2.
FT DOMAIN 663 701 LDL-RECEPTOR CLASS A 3.
FT DOMAIN 723 761 TSP TYPE-1 3.
FT DOMAIN 762 813 TSP TYPE-1 4.
FT DOMAIN 814 867 TSP TYPE-1 4.
FT DISULFID 107 122 BY SIMILARITY.
FT DISULFID 116 127 BY SIMILARITY.
FT DISULFID 129 141 BY SIMILARITY.
FT DISULFID 147 166 BY SIMILARITY.
FT DISULFID 149 169 BY SIMILARITY.
FT DISULFID 171 179 BY SIMILARITY.
FT DISULFID 344 502 BY SIMILARITY.
FT DISULFID 508 520 BY SIMILARITY.
FT DISULFID 515 533 BY SIMILARITY.
FT DISULFID 527 542 BY SIMILARITY.
FT DISULFID 665 677 BY SIMILARITY.
FT DISULFID 672 690 BY SIMILARITY.
FT DISULFID 684 699 BY SIMILARITY.
FT DISULFID 725 737 BY SIMILARITY.
FT DISULFID 732 750 BY SIMILARITY.
FT DISULFID 744 759 BY SIMILARITY.
FT DISULFID 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 409 409 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 867 867
SQ SEQUENCE 867 AA; 91817 MW; 9538F2108E787B49 CRC64;

Query Match
Best Local Similarity 61.0%; Score 36; DB 1; Length 867;
Matches 5; Conservative 83.3%; Pred. No. 23; Mismatches 1; Indels 0; Gaps 0;

Oy 15 CXGPC 20
Db 122 CGGPC 127

RESULT 5
LMB1_HUMAN STANDARD; PRT; 1786 AA.
ID LMB1_HUMAN
AC P07942;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE LAMININ BETA-1 CHAIN PRECURSOR (LAMININ B1 CHAIN).
GN LAMB1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
CC
CC SEQUENCE FROM N.A.
CC MEDLINE=90368768; PubMed=1975589;
CC RA Voelteenano R., Chow L.T., Tryggvason K.;
CC "Structure of the human laminin B1 chain gene.";
CC J. Biol. Chem. 265:15611-15616(1990).
CC [2]
CC RP SEQUENCE FROM N.A.
CC MEDLINE=87280097; PubMed=3611077;
CC RA Plikkarainen T., Eddy R., Fukushima Y., Byers M., Shows T.,
CC RA Plikkarainen T., Saraste M., Tryggvason K.;
CC "Human laminin B1 chain. A multidomain protein with gene (LAMB1)
RT

```



```

FT DOMAIN 1442 1781 COILED COIL (POTENTIAL).
FT DISULFID 271 280 BY SIMILARITY.
FT DISULFID 273 288 BY SIMILARITY.
FT DISULFID 300 309 BY SIMILARITY.
FT DISULFID 312 332 BY SIMILARITY.
FT DISULFID 335 344 BY SIMILARITY.
FT DISULFID 337 362 BY SIMILARITY.
FT DISULFID 365 374 BY SIMILARITY.
FT DISULFID 377 395 BY SIMILARITY.
FT DISULFID 398 411 BY SIMILARITY.
FT DISULFID 400 426 BY SIMILARITY.
FT DISULFID 428 437 BY SIMILARITY.
FT DISULFID 440 455 BY SIMILARITY.
FT DISULFID 458 472 BY SIMILARITY.
FT DISULFID 460 479 BY SIMILARITY.
FT DISULFID 481 490 BY SIMILARITY.
FT DISULFID 493 507 BY SIMILARITY.
FT DISULFID 773 785 BY SIMILARITY.
FT DISULFID 775 792 BY SIMILARITY.
FT DISULFID 794 803 BY SIMILARITY.
FT DISULFID 806 818 BY SIMILARITY.
FT DISULFID 821 833 BY SIMILARITY.
FT DISULFID 823 840 BY SIMILARITY.
FT DISULFID 842 851 BY SIMILARITY.
FT DISULFID 854 864 BY SIMILARITY.
FT DISULFID 867 876 BY SIMILARITY.
FT DISULFID 883 895 BY SIMILARITY.
FT DISULFID 886 914 BY SIMILARITY.
FT DISULFID 898 933 BY SIMILARITY.
FT DISULFID 917 944 BY SIMILARITY.
FT DISULFID 919 944 BY SIMILARITY.

Query Match
Best Local Similarity 61.0%; Score 36; DB 1; Length 1786;
Matches 5; Conservative 83.3%; Pred. No. 39; Indels 0; Gaps 0;
Mismatches 1;

OY 15 CXGPGC 20
Db 1425 CGGPGC 1430

RESULT 6
LMB1_MOUSE STANDARD; PRT; 1786 AA.
AC P02469;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE LAMININ BETA-1 CHAIN PRECURSOR (LAMININ B1 CHAIN).
GN LAMB1-1 OR LAMB-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87147212; PubMed=3493487;
RA Sasaki M., Kato S., Kohno K., Martin G.R., Yamada Y.;
RT "Sequence of the cDNA encoding the laminin B1 chain reveals a
RT multidomain protein containing cysteine-rich repeats."
RL Proc. Natl. Acad. Sci. U.S.A. 84:935-939(1987).
RN [2]
RP SEQUENCE OF 1292-1786 FROM N.A.
RX MEDLINE=85051302; PubMed=6209134;
RA Barlow D.P., Green N.M., Kurkinen M., Hogan B.L.M.;
RT "Sequencing of laminin B chain cDNAs reveals C-terminal regions of
RT coiled-coil alpha-helix."
RL EMBO J. 3:2355-2362(1984).
RN [3]
RP SEQUENCE OF 165-172, 539-547 AND 712-719.
RC STRAIN=BALB/C; TISSUE=Endothelial cells;
RX MEDLINE=97363207; PubMed=9219532;
RA Frieser M., Noeckel H., Pausch F., Roeder C., Hahn A., Deutzmann R.,

```

```

RA Sorokin L.M.;
RT "Cloning of the mouse laminin alpha 4 cDNA. Expression in a subset of
RT endothelium."
RL Eur. J. Biochem. 246:727-735(1997).
CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
CC IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
CC THE BETA-1 CHAIN IS A SUBUNIT OF LAMININ-1 (EHS LAMININ), LAMININ-
CC 2 (MEROSIN), AND LAMININ-6 (K-LAMININ).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR
CC COMPONENT).
CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -1- SIMILARITY: CONTAINS 12.5 LAMININ EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M15525; AAA39407.1; ALT_INIT.
DR EMBL; X05212; CAA28639.1; -.
DR PIR; A26413; MMMSBL.
DR HSSP; P03069; 12IM.
DR MGD; MG1:96743; Lamb1-1.
DR InterPro; IPR000561; -.
DR InterPro; IPR001886; -.
DR InterPro; IPR002049; -.
DR Pfam; PF00053; Laminin_EGF_13.
DR Pfam; PF00055; Laminin_Nterm; 1.
DR PRINTS; PR00011; EGF_LAMININ.
DR PROSITE; PS00022; EGF_1; 9.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01248; LAMININ_TYPE_EGF_11.
DR GlycoProfile; Basement membrane; Extracellular matrix; Coiled coil;
KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
FT SIGNAL 1 21
FT CHAIN 22 1786
FT DOMAIN 22 270
FT DOMAIN 271 540
FT
FT DOMAIN 271 334
FT DOMAIN 335 397
FT DOMAIN 398 457
FT DOMAIN 458 509
FT DOMAIN 510 540
FT DOMAIN 541 772
FT DOMAIN 773 1178
FT
FT DOMAIN 773 820
FT DOMAIN 821 866
FT DOMAIN 867 916
FT DOMAIN 917 975
FT DOMAIN 976 1027
FT DOMAIN 1028 1083
FT DOMAIN 1084 1131
FT DOMAIN 1132 1178
FT DOMAIN 1179 1397
FT DOMAIN 1398 1430
FT DOMAIN 1431 1786
FT DOMAIN 1216 1315
FT DOMAIN 1368 1388
FT DOMAIN 1448 1778
FT DISULFID 271 280

LAMININ EGF-LIKE 1.
LAMININ EGF-LIKE 2.
LAMININ EGF-LIKE 3.
LAMININ EGF-LIKE 4.
LAMININ EGF-LIKE 5 (INCOMPLETE).
LAMININ DOMAIN IV.
8 X LAMININ EGF-LIKE REPEATS (DOMAIN
III).
LAMININ EGF-LIKE 6.
LAMININ EGF-LIKE 7.
LAMININ EGF-LIKE 8.
LAMININ EGF-LIKE 9.
LAMININ EGF-LIKE 10.
LAMININ EGF-LIKE 11.
LAMININ EGF-LIKE 12.
LAMININ EGF-LIKE 13.
DOMAIN II.
DOMAIN ALPHA.
DOMAIN I.
COILED COIL (POTENTIAL).
COILED COIL (POTENTIAL).
COILED COIL (POTENTIAL).
BY SIMILARITY.

```

```

FT DISULFID 273 298 BY SIMILARITY.
FT DISULFID 300 309 BY SIMILARITY.
FT DISULFID 312 332 BY SIMILARITY.
FT DISULFID 335 344 BY SIMILARITY.
FT DISULFID 337 362 BY SIMILARITY.
FT DISULFID 365 374 BY SIMILARITY.
FT DISULFID 377 395 BY SIMILARITY.
FT DISULFID 398 411 BY SIMILARITY.
FT DISULFID 400 426 BY SIMILARITY.
FT DISULFID 428 437 BY SIMILARITY.
FT DISULFID 440 455 BY SIMILARITY.
FT DISULFID 458 472 BY SIMILARITY.
FT DISULFID 460 479 BY SIMILARITY.
FT DISULFID 481 490 BY SIMILARITY.
FT DISULFID 493 507 BY SIMILARITY.
FT DISULFID 773 785 BY SIMILARITY.
FT DISULFID 775 792 BY SIMILARITY.
FT DISULFID 794 803 BY SIMILARITY.
FT DISULFID 806 818 BY SIMILARITY.
FT DISULFID 821 833 BY SIMILARITY.
FT DISULFID 823 840 BY SIMILARITY.
FT DISULFID 842 851 BY SIMILARITY.
FT DISULFID 854 864 BY SIMILARITY.
FT DISULFID 867 876 BY SIMILARITY.
FT DISULFID 883 895 BY SIMILARITY.
FT DISULFID 898 914 BY SIMILARITY.
FT DISULFID 917 933 BY SIMILARITY.
FT DISULFID 919 944 BY SIMILARITY.
FT DISULFID 946 955 BY SIMILARITY.
FT DISULFID 958 973 BY SIMILARITY.
FT DISULFID 976 990 BY SIMILARITY.
FT DISULFID 978 997 BY SIMILARITY.
FT DISULFID 1000 1009 BY SIMILARITY.
FT DISULFID 1012 1025 BY SIMILARITY.
FT DISULFID 1084 1096 BY SIMILARITY.
FT DISULFID 1086 1103 BY SIMILARITY.
FT DISULFID 1105 1114 BY SIMILARITY.
FT DISULFID 1117 1129 BY SIMILARITY.
FT DISULFID 1132 1144 BY SIMILARITY.
FT DISULFID 1134 1151 BY SIMILARITY.
FT DISULFID 1153 1162 BY SIMILARITY.
FT DISULFID 1165 1176 BY SIMILARITY.
FT DISULFID 1179 1179 BY SIMILARITY.
FT DISULFID 1182 1182 INTERCHAIN (PROBABLE).
FT DISULFID 1785 1785 INTERCHAIN (PROBABLE).
FT CARBOHYD 120 120 INTERCHAIN (PROBABLE).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 519 519 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 677 677 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1041 1041 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1195 1195 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1279 1279 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1336 1336 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1343 1343 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1487 1487 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1533 1533 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1542 1542 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1643 1643 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 1531 1534 SCNA -> MEMB (IN REF. 2).
FT CONFLICT 1749 1749 D -> N (IN REF. 2).
SO SEQUENCE 1786 AA; 196904 MW; 846671B7BF41A474 CRC64;

```

Query Match Score 36; DB 1; Length 1786;
 Best Local Similarity 83.3%; Pred. No. 39;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 CXGPC 20
 DB 1425 CCGPGC 1430

```

RESULT 7
ID Y085_METUA STANDARD: PRT; 373 AA.
AC O57550:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL PROTEIN MJ0085.
GN MJ0085.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN 111
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geophagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii".
RL Science 273:1058-1073(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U67466; AAB98067.1; -.
DR InterPro: IPR002491; -.
DR Pfam: PF01497; Peripla_BP_2; 1.
DR Hypothetical protein.
KW SEQUENCE 373 AA; 41812 MW; 8B64F5683E830878 CRC64;

```

Query Match Score 35; DB 1; Length 373;
 Best Local Similarity 83.3%; Pred. No. 18;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 CXGPC 20
 DB 61 CCGPGC 66

```

RESULT 8
ID ATF2_CHICK STANDARD: PRT; 487 AA.
AC O93602:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CYCLIC-AMP-DEPENDENT TRANSCRIPTION FACTOR ATF-2 (ACTIVATING
DE TRANSCRIPTION FACTOR 2).
DE ATF2.
GN Gallus gallus (Chicken).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN 111
RP SEQUENCE FROM N.A.
RA Hugnier S., Baguet J., Perez S., van Dam H., Castellazzi M.;
RT "Transcription factor ATF2 cooperates with v-Jun to promote growth-

```

RT factor independent proliferation in vitro and tumor formation in
 RT vivo.";
 RT Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: A TRANSCRIPTIONAL ACTIVATOR, PROBABLY CONSTITUTIVE,
 CC WHICH BINDS TO THE CAMP RESPONSIVE REGION (CRE) (CONSENSUS:
 CC 5'GTGACGT(A/C)(A/G)-3'); A SEQUENCE PRESENT IN MANY VIRAL AND
 CC CELLULAR PROMOTERS (BY SIMILARITY).
 CC -1- SUBUNIT: BINDS DNA AS A DIMER.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- SIMILARITY: BELONGS TO THE BZIP FAMILY. ATF SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y17724; CAA76838.1; -.
 DR HSSP: P08047; ISP2.
 DR InterPro: IPR000822; -.
 DR InterPro: IPR001871; -.
 DR Pfam: PF001170; bzip: 1.
 DR Pfam: PF000096; zf-C2H2; 1.
 DR PROSITE: PS00036; BZIP_BASIC; 1.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 1.
 DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 1.
 KM Transcription regulation: DNA-binding; Activator; Nuclear protein;
 KM Zinc-finger; Metal-binding.
 FT ZN_FING 7 31 C2H2-TYPE.
 FT DNA_BIND 333 356 BASIC MOTIF.
 FT DOMAIN 362 390 LEUCINE-ZIPPER.
 SQ SEQUENCE 487 AA; 52406 MW; A1FA2734D9C6A146 CRC64;

Query Match 52.5%; Score 31; DB 1; Length 487;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 15 CXGPGC 20
 Db 9 CTAPGC 14

RESULT 9
 ATF2_HUMAN
 ID ATF2_HUMAN STANDARD; PRT; 487 AA.
 AC P15336;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE CYCLIC-AMP-DEPENDENT TRANSCRIPTION FACTOR ATF-2 (ACTIVATING
 DE TRANSCRIPTION FACTOR 2) (CAMP RESPONSE ELEMENT BINDING PROTEIN CRE-
 DE Bp1) (HB16).
 GN ATF2 OR CREB2 OR CREBp1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCB1_TaxID=9606;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=90005408; PubMed=2529117;
 RA Maekawa T., Sakura H., Kanel-Tsili C., Sudo T., Yoshimura T.,
 RA Maekawa J.I., Yoshida M., Ishii S.;
 RT Leucine zipper structure of the protein CRE-Bp1 binding to the
 RT cyclic AMP response element in brain.";
 RL EMBO J. 8:2023-2028(1989).
 RN [2]
 RP SEQUENCE OF 193-487 FROM N.A.
 RA MEDLINE=90205810; PubMed=2320002;
 RA Kara C.J., Liou H.-C., Iwashiki L.B., Glimcher L.H.;

RT "A cDNA for a human cyclic AMP response element-binding protein which
 RT is distinct from CREB and expressed preferentially in brain.";
 RL Mol. Cell. Biol. 10:1347-1357(1990).
 CC -1- FUNCTION: THIS PROTEIN BINDS THE CAMP RESPONSE ELEMENT (CRE)
 CC (CONSENSUS: 5'GTGACGT(A/C)(A/G)-3'), A SEQUENCE PRESENT IN MANY
 CC VIRAL AND CELLULAR PROMOTERS.
 CC -1- SUBUNIT: BINDS DNA AS A DIMER.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- ALTERNATIVE PRODUCTS: MBP AND CRE-BP1 MAY BE MEMBERS OF A FAMILY
 CC OF MBP/CRE-BP PROTEINS GENERATED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: ABUNDANT EXPRESSION SEEN IN THE BRAIN.
 CC -1- PM: PHOSPHORYLATION OF THR-69 AND THR-71 BY MAPK14 CAUSES
 CC INCREASED TRANSCRIPTIONAL ACTIVITY. ALSO PHOSPHORYLATED AND
 CC ACTIVATED BY JNK.
 CC -1- SIMILARITY: BELONGS TO THE BZIP FAMILY. ATF SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X15875; CAA33886.1; -.
 DR EMBL: M31630; AAA35951.1; -.
 DR PIR: S05380; S05380.
 DR HSSP: P08047; ISP2.
 DR TRANSFAC: T00167; -.
 DR MIM: 123811; -.
 DR InterPro: IPR000822; -.
 DR InterPro: IPR001871; -.
 DR Pfam: PF001170; bzip: 1.
 DR Pfam: PF000096; zf-C2H2; 1.
 DR PROSITE: PS00036; BZIP_BASIC; 1.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 1.
 DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 1.
 KM Transcription regulation: DNA-binding; Activator; Phosphorylation;
 KM Nuclear protein; Alternative splicing; Zinc-finger; Metal-binding.
 FT ZN_FING 7 31 C2H2-TYPE.
 FT DNA_BIND 333 356 BASIC MOTIF.
 FT DOMAIN 362 390 LEUCINE-ZIPPER.
 FT MOD_RES 51 51 PHOSPHORYLATION (BY MAPK14).
 FT MOD_RES 53 53 PHOSPHORYLATION (BY MAPK14).
 SQ SEQUENCE 487 AA; 52249 MW; EEE4D17EE8DB7C07 CRC64;

Query Match 52.5%; Score 31; DB 1; Length 487;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 15 CXGPGC 20
 Db 9 CTAPGC 14

RESULT 10
 ATF2_MOUSE
 ID ATF2_MOUSE STANDARD; PRT; 487 AA.
 AC P16951; Q64089; Q64090; Q64091;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE CYCLIC-AMP-DEPENDENT TRANSCRIPTION FACTOR ATF-2 (ACTIVATING
 DE TRANSCRIPTION FACTOR 2) (CAMP RESPONSE ELEMENT BINDING PROTEIN CRE-
 DE Bp1) (MBP PROTEIN).
 GN ATF2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCB1_TaxID=10090;
 OX [1]
 RP SEQUENCE OF 9-487 FROM N.A., AND ALTERNATIVE SPLICING.

RA MEDLINE:92123199; PubMed=1531087;
RX Georgopoulos K., Morgan B.A., Moore D.D.;
RT "Functionally distinct isoforms of the CRE-BP DNA-binding protein
RL mediate activity of a T-cell-specific enhancer.";
RN Mol. Cell. Biol. 12:747-757(1992).
[2]
RP SEQUENCE OF 77-487 FROM N.A.
RX MEDLINE=90205841; PubMed=2138707;
RA Ivashchyk L.B., Liou H.-C., Kara C.J., Lamph W.W., Verma I.M.,
RT "MBP/CRE-BP2 and c-Jun form a complex which binds to the cyclic AMP,
RL element.";
RN Mol. Cell. Biol. 10:1609-1621(1990).
CC -1- FUNCTION: A TRANSCRIPTIONAL ACTIVATOR, PROBABLY CONSTITUTIVE,
CC WHICH BINDS TO THE CAMP RESPONSIVE REGION (CRE) (CONSENSUS:
CC 5'-GTGACG(A/C)(A/G)-3'); A SEQUENCE PRESENT IN MANY VIRAL AND
CC CELLULAR PROMOTERS. THE INTERACTION OF MBP/CRE-BP2 WITH C-JUN
CC REDIRECTS C-JUN TO BIND TO CREs PREFERENTIALLY OVER THE 12-O-
CC TERADECANOYLPHORBOL-13-ACETATE RESPONSE ELEMENTS (TREs) AS PART
CC OF AN MBP-C-JUN COMPLEX.
CC -1- SUBUNIT: BINDS DNA AS A DIMER AND CAN FORM A HOMODIMER IN THE
CC ABSENCE OF DNA. CAN FORM AN HETERODIMER WITH C-JUN.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: 1 (SHOWN HERE), 2 AND 3; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: BELONGS TO THE BP1 FAMILY. AMP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S76657; AAB21128.1; ALT_INT.
DR EMBL: S76659; AAB21129.1; ALT_INT.
DR EMBL: S76655; AAB21127.1; -
DR EMBL: M31629; AAA39780.1; -
DR PIR: A34785; A34785.
DR HSSP: P03069; PDGC.
DR TRASNSEC: T01017; -
DR MGD: MGI:109349; Atf2.
DR InterPro: IPR000822; -
DR InterPro: IPR001871; -
DR Pfam: PF00170; bZIP; 1.
DR Pfam: PF00096; zf-C2H2; 1.
DR PROSITE: PS00036; BP1_BASIC; 1.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE: PS00157; ZINC_FINGER_C2H2_2; 1.
DR Transcription regulation; DNA-binding; Activator; Phosphorylation;
KW Nuclear protein; Alternative splicing; Zinc-finger; Metal-binding.
FT ZN_FING 7 31
FT DNA_BIND 333 356
FT DOMAIN 362 390
FT MOD_RES 51 51
FT MOD_RES 53 53
FT MOD_RES 53 53
FT VARSPLIC 1 48
FT VARSPLIC 1 48
FT VARSPLIC 132 229
FT CONFLICT 482 487
FT CONFLICT 487 AA; 52297 MW; PSCDC3BC3119ACB CRC64;
SQ SEQUENCE 487 AA; 52297 MW; PSCDC3BC3119ACB CRC64;

Query Match	52.5%;	Score 31;	DB 1;	Length 487;
Best Local Similarity	66.7%;	Pred. NO. 1.1e+02;		
Matches 4;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

DB	9	CTAPC	14
RESULT	11		
ID	ATF2_RAT	STANDARD:	PRT: 487 AA.
AC	000969; 062870;		
DT	01-APR-1993 (Rel. 25, Created)		
DT	30-MAY-2000 (Rel. 39, Last sequence update)		
DT	30-MAY-2000 (Rel. 39, Last annotation update)		
DE	CYCLOIC-AMF-DEPENDENT TRANSCRIPTION FACTOR ATF-2 (ACTIVATING		
DE	TRANSCRIPTION FACTOR 2) (CAMP RESPONSE ELEMENT BINDING PROTEIN CRE-		
DE	BPI).		
GN	ATF2.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A. (ISOFORM 1).		
RA	STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;		
RC	Muramatsu S.;		
RL	Submitted (OCT-1995) to the EMBL/GenBank/DDBJ databases.		
RL	[2]		
RP	SEQUENCE FROM N.A. (ISOFORM 2).		
RC	TISSUE=Brain;		
RX	MEDLINE=91332085; PubMed=1714459;		
RA	Kageyama R., Sasai Y., Nakanishi S.;		
RT	"Molecular characterization of transcription factors that bind to the		
RT	CAMP responsive region of the substance P precursor gene. cDNA		
RT	cloning of a novel C/EBP-related factor.";		
RL	J. Biol. Chem. 266:15525-15531(1991).		
CC	-1- FUNCTION: A TRANSCRIPTIONAL ACTIVATOR, PROBABLY CONSTITUTIVE,		
CC	WHICH BINDS TO THE CAMP RESPONSIVE REGION (CRE) (CONSENSUS:		
CC	5'GTGCGT(A/C)(A/G)-3'); A SEQUENCE PRESENT IN MANY VIRAL AND		
CC	CELLULAR PROMOTERS.		
CC	-1- SUBUNIT: BINDS DNA AS A DIMER.		
CC	-1- SUBCELLULAR LOCATION: NUCLEAR.		
CC	-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE		
CC	PRODUCED BY ALTERNATIVE SPLICING.		
CC	-1- SIMILARITY: BELONGS TO THE BZIP FAMILY. ATP SUBFAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; U38938; AAA93263.1; -;		
DR	EMBL; M65148; AAA42013.1; -;		
DR	PIR; A39429; A39429.		
DR	HSSP; P08047; 1SP2.		
DR	TRANSFAC; T01382; -;		
DR	InterPro; IPR000822; -;		
DR	InterPro; IPR001871; -;		
DR	pfam; PF001070; bZIP; 1.		
DR	pfam; pf000066; zf-C2H2; 1		
DR	PROSITE; PS00036; bZIP_BASIC; 1.		
DR	PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.		
DR	PROSITE; PS50157; ZINC_FINDER_C2H2_2; 1.		
KW	Transcription regulation; DNA-binding; Activator; Phosphorylation;		
KW	Nuclear protein; Alternative splicing; Zinc-finger; Metal-binding.		
FT	ZN_FING	7	31
FT	DNA_BIND	333	356
FT	DOMAIN	362	390
FT	MOD_RES	51	51
FT	MOD_RES	53	53
FT	MOD_RES	53	53
FT	VANSPLIC	132	229
SO	SEQUENCE	487 AA; 52286 MM;	MISSING (IN ISOFORM 2). 4BD95B10DFF9EE CRC64;

Query Match 52.5%; Score 31; DB 1; Length 487;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 15 CXGPC 20
 1 111
 Db 9 CTAAPC 14

RESULT 12

POLY_RUBV STANDARD; PRT; 522 AA.
 AC P08564; Q88780;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE STRUCTURAL POLYPROTEIN [CONTAINS: SPIKE GLYCOPROTEINS E1 AND E2]
 DE (FRAGMENT).
 OS Rubella virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
 OC Rubivirus.
 OX NCBI_TaxID=11041;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87057359; PubMed=3023358;
 RA Nakhasi H.L., Meyer B.C., Liu T.Y.;
 RT "Rubella virus cDNA. Sequence and expression of E1 envelope protein.";
 RL J. Biol. Chem. 261:16616-16621(1986)
 CC -1- FUNCTION: GLYCOPROTEIN E1 CONTAINS THE VIRAL HEMAGGLUTININ
 CC ACTIVITY.
 CC -1- SUBUNIT: MULTIPLE COPIES OF THE C PROTEIN COMPRISE THE
 CC NUCLEOCAPSID.
 CC -1- DISEASE: RUBELLA INFECTION DURING THE FIRST TRIMESTER OF PREGNANCY
 CC RESULTS IN DEATH OR MULTISYSTEM BIRTH DEFECTS & MICROCEPHALY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: J02620; AAA47423.1; -;
 DR PIR: A25340; GNMVRL.
 DR Polyprotein; Glycoprotein; Transmembrane; Signal.
 KW NON_TER 1
 FT CHAIN 1
 FT SIGNAL <1 92 SPIKE GLYCOPROTEIN E2.
 FT CHAIN 93 112
 FT CHAIN 113 522 SPIKE GLYCOPROTEIN E1.
 FT CARBOHD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 522 AA; 55958 MW; FDOB90BC202DA8BD CRC64;

Query Match 52.5%; Score 31; DB 1; Length 522;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 15 CXGPC 20
 1 111
 Db 120 CTAAPC 125

RESULT 13
 YFHE_YEAST STANDARD; PRT; 535 AA.
 AC P43590;
 DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE HYPOTHETICAL 61.8 KDA PEPTIDASE IN MPRI-GCN20 INTERGENIC REGION
 DE (EC 3.4.-.-).
 GN YFR006W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX MEDLINE=95400292; PubMed=7670463;
 RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
 RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
 RA Yamazaki M., Tashiro H., Eki T.;
 RT "Analysis of the nucleotide sequence of chromosome VI from
 RT Saccharomyces cerevisiae.";
 RL Nat. Genet. 10:261-268(1995).

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M24B.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL: D50617; BAA09245.1; -;
 DR MEROPS: M24.007; -;
 DR SCD: S0001902; YFR006W.
 DR InterPro: IPR000994; -;
 DR InterPro: IPR001131; -;
 DR Pfam: PF00557; Peptidase_M24; 1.
 DR PROSITE: PS00491; PROLINE PEPTIDASE; 1.
 KW Hypothetical protein; Hydrolase; Transmembrane.
 FT TRANSLEM 8 24 POTENTIAL.
 SQ SEQUENCE 535 AA; 61753 MW; 4D251041CE9627FC CRC64;

Query Match 52.5%; Score 31; DB 1; Length 535;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 15 CXGPC 20
 1 111
 Db 290 CTAAPC 295

RESULT 14
 HS7H_HUMAN STANDARD; PRT; 641 AA.
 AC P34931; O9UOM1;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE HEAT SHOCK 70 KDA PROTEIN 1-HOM (HSP70-HOM).
 GN HSPAL1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91055806; PubMed=1700760;
 RA Milner C.M., Campbell R.D.;
 RT "Structure and expression of the three MHC-linked HSP70 genes.";
 RL Immunogenetics 32:242-251(1990).
 [2]

RP SEQUENCE FROM N.A.
 RP Rowen L., Qin S., Madan A., Dickhoff R., Dors M., Madan A., Hicks P.,
 RA Loretz C., Ratcliffe A., Abbasi N., Shaffer T., Hood L.;

RT "Sequence of the human major histocompatibility complex class III
region."
RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
RN [3]
RP VARIANT MET-493.
RX MEDLINE-92406261; PubMed-1356099;
RA Milner C.M., Campbell R.D.;
RT "Polymorphic analysis of the three MHC-linked HSP70 genes."
RL Immunogenetics 36:357-362(1992).
CC -1- INDUCTION: NOT INDUCED BY HEAT SHOCK.
CC -1- MISCELLANEOUS: ENCODED IN THE MHC-III COMPLEX.
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M59829; AAA63228.1; -
DR EMBL: AF134726; AAD21817.1; -
DR PIR: B45871; B45871.
DR HSSP: P19120; INGT.
DR MIM: 140559; -
DR InterPro: IPR001023; -
DR Pfam: PF00012; HSP70; 1.
DR PRINTS: PR00301; HEATSHOCK70.
DR PROSITE: PS00297; HSP70_1; 1.
DR PROSITE: PS00329; HSP70_2; 1.
DR PROSITE: PS01036; HSP70_3; 1.
KW ATP-binding; Multigene family; MHC III; Polymorphism.
FT VARIANT 493 493
T->M.
/FTid=VAR_003820.
FT CONFLICT 408 408 V->A (IN REF. 2).
FT CONFLICT 424 424 P->T (IN REF. 2).
SQ SEQUENCE 641 AA; 70399 MW; 5CE758C9CF6E55 CRC64;

Query Match 52.5%; Score 31; DB 1; Length 641;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 15 CXGPGC 20
Db 617 CTGPAC 622

RESULT 15
POLY_RUBVM STANDARD; PRT; 992 AA.
AC P08563;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE STRUCTURAL POLYPEPTIDE [CONTAINS: NUCLEOCAPSID PROTEIN C; MEMBRANE
DE GLYCOPROTEIN E1 AND E2].
OS Rubella virus (strain M33).
OC Viruses: ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Rubivirus.
OX NCBI_Taxid-11043;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-87174825; PubMed-3562245;
RA Clarke D.M., Loo T.W., Hui I., Chong P., Gillam S.;
RT "Nucleotide sequence and in vitro expression of rubella virus 24S
RT subgenomic messenger RNA encoding the structural proteins E1, E2 and
RT C."
RL Nucleic Acids Res. 15:3041-3057(1987).
CC -1- FUNCTION: GLYCOPROTEIN E1 CONTAINS THE VIRAL HEMAGGLUTININ
CC ACTIVITY.
CC -1- SUBUNIT: MULTIPLE COPIES OF THE C PROTEIN COMPRISE THE

CC NUCLEOCAPSID.
CC -1- DISEASE: RUBELLA INFECTION DURING THE FIRST TRIMESTER OF PREGNANCY
CC RESULTS IN DEATH OR MULTISYSTEM BIRTH DEFECTS & MICROCEPHALY
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X05259; CAA28880.1; -
DR PIR: A27505; GNMWR3.
KW Polypeptide; Nucleocapsid; Glycoprotein; Transmembrane; Signal.
FT CHAIN 1 299
FT SIGNAL 300 559
FT CHAIN 560 580
FT CHAIN 581 992
FT TRANSMEM 277 297
FT TRANSMEM 515 531
FT TRANSMEM 533 553
FT CARBOHYD 352 352
FT CARBOHYD 370 370
FT CARBOHYD 656 656
FT CARBOHYD 757 757
FT CARBOHYD 789 789
SQ SEQUENCE 992 AA; 106905 MW; D268889C9E78CF1E CRC64;

Query Match 52.5%; Score 31; DB 1; Length 992;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 15 CXGPGC 20
Db 588 CTAPGC 593

Search completed: November 21, 2001, 15:46:25
Job time: 88 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 15:44:57 ; Search time 22.94 Seconds

(Without alignments)
66.412 Million cell updates/sec

Title: US-09-609-721-14

Perfect score: 122

Sequence: 1 QVRESWCTGPGCACLQACL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	46.7	156	2 B83361	probable transcrip
2	51	41.8	148	2 B82503	soxr protein VCA00
3	50	41.0	72	2 S39416	metallothionein 10
4	50	41.0	72	2 S39418	metallothionein 10
5	50	41.0	72	2 S39419	metallothionein 10
6	50	41.0	72	2 S39417	metallothionein 10
7	50	41.0	220	2 A43540	mb-1 protein precu
8	50	41.0	659	2 H69421	heterodisulfide re
9	49	40.2	37	2 E44007	apoptoxin III - tr
10	48	39.3	67	2 F81037	hypothetical prote
11	48	39.3	71	2 D81983	hypothetical prote
12	48	39.3	518	1 PRBE11	56k serine protein
13	48	39.3	956	1 RRBWSC	probable RNA-direc
14	47.5	38.9	161	2 S55608	hypothetical prote
15	47	38.5	385	2 B85850	probable transport
16	47	38.5	385	2 A64941	yeih protein - Esc
17	46.5	38.1	315	2 T21141	hypothetical prote
18	46	37.7	411	2 T39481	hypothetical prote
19	46	37.7	544	2 T40752	hypothetical prote
20	46	37.7	986	1 OYURCP	speract receptor p
21	46	37.7	1125	1 OYURCP	speract receptor p
22	45.5	37.3	61	2 A37425	metallothionein 2
23	45.5	37.3	62	2 S54335	metallothionein-2c
24	45.5	37.3	62	2 S54336	metallothionein-2a
25	45.5	37.3	387	2 E83679	multidrug-efflux t
26	45	36.9	175	2 T36798	probable transcrip
27	45	36.9	248	2 T40343	hypothetical ubiqn
28	45	36.9	324	2 F22848	hypothetical ORF-4
29	45	36.9	344	2 C25877	hypothetical prote

30	45	36.9	1136	2 A56559	enhancer-trap-loou
31	45	36.9	1255	1 A24571	protein-tyrosine k
32	44.5	36.5	47	2 S11529	gamma-purothionin
33	44.5	36.5	47	2 S11530	purothionin gamma
34	44.5	36.5	47	2 S13849	hordothionin gamma
35	44.5	36.5	180	2 E69269	molybdopterin oxid
36	44.5	36.5	831	2 T15108	hypothetical prote
37	44.5	36.5	1420	2 A32869	apolipoprotein(a)
38	44	36.1	128	2 T15101	hypothetical prote
39	44	36.1	242	2 S41995	tryptophan synthas
40	44	36.1	242	2 T14365	probable tryptopha
41	44	36.1	340	2 S54821	chrp protein - Erw
42	44	36.1	433	2 H96549	hypothetical prote
43	44	36.1	732	1 S07624	acylaminoacyl-pept
44	44	36.1	790	2 E86460	hypothetical prote
45	44	36.1	1088	2 T14917	homeotic protein p

ALIGNMENTS

```
RESULT 1
B83361
Probable transcription regulator PA2273 [Imported] - Pseudomonas aeruginosa (strain P
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: B83361
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kae, A.; Lartig, K.; L
.: Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337
A:Accession: B83361
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-156 <STO>
A:Cross-references: GB:AE004653; GB:AE004091; NID:g9948298; PIDN:AAG05661.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA2273

Query Match 46.7%; Score 57; DB 2; Length 156;
Best Local Similarity 76.9%; Pred. No. 1.1;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 7 GCIGPGCACLQAC 19
Db 116 GCIGGCCLSLQAC 128

RESULT 2
B82503
soxr protein VCA0084 [Imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: B82503
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.
Charlson, D.; Ermolaeva, M.D.; Vamthyan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers
I, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae
A:Reference number: A82035; MUID:20406833
A:Accession: B82503
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-148 <HEI>
A:Cross-references: GB:AE004351; GB:AE003853; NID:g9657462; PIDN:AAF95998.1; GSPDB:GN
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCA0084
A:Map position: 2
```


Query Match 41.8%; Score 51; DB 2; Length 148;
Best Local Similarity 61.5%; Pred. No. 6.3;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 7 GCIGPGCACLQAC 19
||| | | | |
Db 109 GCGGCGRCGDCAC 121

RESULT 3

S39416
metallothionein 10-I - blue mussel
C:Species: Mytilus edulis (blue mussel)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Mar-1999
C:Accession: S39416
R:MacKay, E.A.; Overnell, J.; Dunbar, B.; Davidson, I.; Hunziker, P.E.; Kaegi, J.H.R.; F
Eur. J. Biochem. 218, 183-194, 1993
A:Title: Complete amino acid sequences of five dimeric and four monomeric forms of metal
A:Reference number: S39416; MUID:94062828
A:Accession: S39416
A:Molecule type: protein
A:Residues: 1-72 <MAC>
C:Superfamily: metallothionein
C:Keywords: metal binding

Query Match 41.0%; Score 50; DB 2; Length 72;
Best Local Similarity 61.5%; Pred. No. 5;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 7 GCIGPGCACLQAC 19
||| | | | |
Db 17 GCGGCGRCGDCAC 29

RESULT 4

S39418
metallothionein 10-III - blue mussel
C:Species: Mytilus edulis (blue mussel)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Mar-1999
C:Accession: S39418
R:MacKay, E.A.; Overnell, J.; Dunbar, B.; Davidson, I.; Hunziker, P.E.; Kaegi, J.H.R.; F
Eur. J. Biochem. 218, 183-194, 1993
A:Title: Complete amino acid sequences of five dimeric and four monomeric forms of metal
A:Reference number: S39416; MUID:94062828
A:Accession: S39418
A:Molecule type: protein
A:Residues: 1-72 <MAC>
C:Superfamily: metallothionein
C:Keywords: metal binding

Query Match 41.0%; Score 50; DB 2; Length 72;
Best Local Similarity 61.5%; Pred. No. 5;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 7 GCIGPGCACLQAC 19
||| | | | |
Db 17 GCGGCGRCGDCAC 29

RESULT 5

S39419
metallothionein 10-IV - blue mussel
C:Species: Mytilus edulis (blue mussel)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Mar-1999
C:Accession: S39419
R:MacKay, E.A.; Overnell, J.; Dunbar, B.; Davidson, I.; Hunziker, P.E.; Kaegi, J.H.R.; F
Eur. J. Biochem. 218, 183-194, 1993
A:Title: Complete amino acid sequences of five dimeric and four monomeric forms of metal
A:Reference number: S39416; MUID:94062828

A:Accession: S39419
A:Molecule type: protein
A:Residues: 1-72 <MAC>
C:Superfamily: metallothionein
C:Keywords: metal binding

Query Match 41.0%; Score 50; DB 2; Length 72;
Best Local Similarity 61.5%; Pred. No. 5;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 7 GCIGPGCACLQAC 19
||| | | | |
Db 17 GCGGCGRCGDCAC 29

RESULT 6

S39417
metallothionein 10-II - blue mussel
C:Species: Mytilus edulis (blue mussel)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Mar-1999
C:Accession: S39417
R:MacKay, E.A.; Overnell, J.; Dunbar, B.; Davidson, I.; Hunziker, P.E.; Kaegi, J.H.R.
Eur. J. Biochem. 218, 183-194, 1993
A:Title: Complete amino acid sequences of five dimeric and four monomeric forms of me
A:Reference number: S39416; MUID:94062828
A:Accession: S39417
A:Molecule type: protein
A:Residues: 1-72 <MAC>
C:Superfamily: metallothionein
C:Keywords: metal binding

Query Match 41.0%; Score 50; DB 2; Length 72;
Best Local Similarity 61.5%; Pred. No. 5;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 7 GCIGPGCACLQAC 19
||| | | | |
Db 17 GCGGCGRCGDCAC 29

RESULT 7

A43540
mb-1 protein precursor - mouse
N:Alternate names: surface Igm complex alpha chain
C:Species: Mus musculus (house mouse)
C>Date: 06-Nov-1992 #sequence_revision 06-Nov-1992 #text_change 05-Nov-1999
C:Accession: A43540; S01648; A60228; A39398; I57521
R:Kashiwamura, S.I.; Koyama, T.; Matsuo, T.; Steinmetz, M.; Kimoto, M.; Sakaguchi, N.
J. Immunol. 145, 337-343, 1990
A:Title: Structure of the murine mb-1 gene encoding a putative sigM-associated molecu
A:Reference number: A43540; MUID:90293481
A:Accession: A43540

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-220 <KAS>
A:Cross-references: GB:M31773; NID:q199032; PIDN:AAA39494.1; PID:9387413
R:Sakaguchi, N.; Kashiwamura, S.; Kimoto, M.; Thalmann, P.; Melchers, F.
EMBO J. 7, 3457-3464, 1988
A:Title: B lymphocyte lineage-restricted expression of mb-1, a gene with CD3-like str
A:Reference number: S01648; MUID:89091088
A:Accession: S01648
A:Molecule type: mRNA
A:Residues: 1-220 <SAK>
A:Cross-references: EMBL:X13450; NID:g53015; PIDN:CAA31801.1; PID:g53016
R:Hombach, J.; Lottspeich, F.; Reith, W.
Eur. J. Immunol. 20, 2795-2799, 1990
A:Title: Identification of the genes encoding the Igm-alpha and Ig-beta components of
A:Reference number: A60228; MUID:91099432
A:Accession: A60228
A:Molecule type: protein
A:Residues: 29-38 <HOM>

Campbell, K.S.; Hager, E.J.; Friedrich, R.J.; Cambler, J.C.
Proc. Natl. Acad. Sci. U.S.A. 88, 3982-3986, 1991

A>Title: Igm antigen receptor complex contains phosphoprotein products of B29 and mb-1 genes
A:Reference number: A39398; MUID:91219496
A:Accession: A39398

A:Molecule type: protein
A:Residues: 'X', 30-38 <CAM>

R:Travis, A.; Hagman, J.; Grosschedl, R.
Mol. Cell. Biol. 11, 5756-5766, 1991

A>Title: Heterogeneously initiated transcription from the pre-B- and B-cell-specific mb-1 promoter
A:Reference number: 157521; MUID:92017857
A:Accession: 157521

A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-15, 'QA' <RES>
A:Cross-references: GB:S59359; NID:g237080
C:Genetics:

A:Gene: mb-1
C:Keywords: membrane protein
F:1-28/Domain: signal sequence #status predicted <STG>
F:29-220/Product: mb-1 protein #status predicted <MAT>

Query Match 41.0%; Score 50; DB 2; Length 220;
Best Local Similarity 43.8%; Pred. No. 12;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 VYESWGCIGPGACACLO 17
::: ||| :
Db 15 LFLSYACLGGCCOALR 30

RESULT 8
H69421 heterodisulfide reductase, subunit A (hdra-2) homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 17-Mar-2000
C:Accession: H69421

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997

A:Authors: Ulteback, T.; Cotton, M.D.; Spriggs, T.; Artlach, P.; Kalne, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.

A>Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon Pyrococcus furiosus
A:Reference number: A69250; MUID:98049343
A:Accession: H69421

A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-659 <KL>

A:Cross-references: GB:AE001009; GB:AE00782; NID:g2689332; PIDN:AAB9867.1; PID:g2649115
C:Superfamily: unassigned ferredoxin [2(Fe-4S)]-related proteins; ferredoxin [2(Fe-4S)] hdb

Query Match 41.0%; Score 50; DB 2; Length 659;
Best Local Similarity 55.0%; Pred. No. 26;
Matches 11; Conservative 2; Mismatches 5; Indels 2; Gaps 2;

OY 1 QVESWG-CIGPCACLOAC 19
||| ||| |::||
Db 232 QTYVDMDKICGCG-ACYEAC 250

RESULT 9
E44007 aptotoxin III - trap-door spider (Aptostichus sp.)
N:Alternate names: insecticidal peptide Aps III
C:Species: Aptostichus sp.
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: E44007
R:Skinner, W.S.; Dennis, P.A.; Li, J.P.; Quistad, G.B.
Toxicom 30, 1043-1050, 1992

A>Title: Identification of insecticidal peptides from venom of the trap-door spider, Aptostichus

A:Reference number: A44007; MUID:93069259
A:Accession: E44007
A:Molecule type: protein
A:Residues: 1-57 <SKT>
A:Cross-references: PIDN:AAB24051.1; PID:g259281
A>Note: the source is designated as Aptostichus schlinger1
A>Note: sequence extracted from NCBI backbone (NCBIP:119526)
C:Keywords: disulfide bond; toxin; venom

Query Match 40.2%; Score 49; DB 2; Length 37;
Best Local Similarity 58.3%; Pred. No. 4.1;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 3 YESMCGICPGCA 14
 | | | | |
Db 21 YNWMCIGGCS 32

RESULT 10
F81037
hypothetical protein NMB1825 [imported] - Neisseria meningitidis (strain MC58 serogro
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: F81037
R:Retelien, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
ti, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignanl, V.; Pizze, M.
Science 287, 1809-1815, 2000
A:Authors: Grandt, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappunli, R.;
A>Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: AB1000; MUID:20175755
A:Accession: F81037
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-67 <TEU>
A:Cross-references: GB:AE002532; GB:AE002098; NID:g7227078; PIDN:AAF42160.1; PID:g7222
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1825

Query Match 39.3%; Score 48; DB 2; Length 67;
Best Local Similarity 56.2%; Pred. No. 8.6;
Matches 9; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

OY 4 ESMGCIGPGC--ACLO 17
 ||| | | | | | | | |
Db 25 ESMGCCSDCGEACIO 40

RESULT 11
D81983
hypothetical protein NMA0634 [imported] - Neisseria meningitidis (strain Z2491 serogr
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: D81983
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo
R: Holtroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre
Nature 404, 502-506, 2000
A>Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491
A:Reference number: AB1775; MUID:20222556
A:Accession: D81983
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-71 <PAP>
A:Cross-references: GB:AL162753; GB:AL157959; NID:g7379120; PIDN:CAB83924.1; PID:g737
C:Genetics:
A:Experimental source: serogroup A, strain Z2491
A:Gene: NMA0634

Query Match 39.3%; Score 48; DB 2; Length 71;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 15:44:57 ; Search time 12.78 Seconds

(without alignments)
53.608 Million cell updates/sec

Title: US-09-609-721-14

Perfect score: 122

Sequence: 1 QVIESMGCIQPGCACLQACL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Database : SwissProt_39:*

Result No.	Score	Query Match	Length	DB ID	Description
1	57	46.7	156	1	SOXR_PSEAE
2	50	41.0	72	1	MT11_MTTED
3	50	41.0	72	1	MT12_MTTED
4	50	41.0	72	1	MT13_MTTED
5	50	41.0	72	1	MT14_MTTED
6	50	41.0	220	1	C79A_MOUSE
7	49	40.2	37	1	TXP3_APTSC
8	48	39.3	518	1	VG47_HSV1
9	48	39.3	956	1	RPOD_SBMV
10	47	38.5	385	1	YEHY_ECOLI
11	46	37.7	470	1	PROP_CAVPO
12	46	37.7	986	1	CYGR_ARBPV
13	46	37.7	1125	1	CYGS_SRRPV
14	45.5	37.3	62	1	MT2A_RABIT
15	45.5	37.3	62	1	MT2C_RABIT
16	45	36.9	1255	1	ERB2_HUMAN
17	44.5	36.5	47	1	THG1_WHEAT
18	44.5	36.5	47	1	THG2_WHEAT
19	44.5	36.5	47	1	THG3_WHEAT
20	44.5	36.5	1420	1	APOR_MACMU
21	44	36.1	242	1	TRPA_CYACA
22	44	36.1	340	1	CBBR_ERMCH
23	44	36.1	614	1	NTBE_RABIT
24	44	36.1	732	1	ACPH_RAT
25	44	36.1	1088	1	PRH_PPCRC
26	43.5	35.7	1135	1	VGJM_HANTL
27	43.5	35.7	1135	1	VGJM_HANTL
28	43	35.2	395	1	CRIC_CAEEL
29	42.5	34.8	60	1	FER_MERTL
30	42.5	34.8	74	1	MT_CRAVI
31	42.5	34.8	641	1	HS73_RAT
32	42.5	34.8	1210	1	EGFR_MOUSE
33	42	34.4	429	1	YNBH_YEAST

ALIGNMENTS

```

RESULT 1
SOXR_PSEAE STANDARD: PRT: 156 AA.
ID SOXR_PSEAE
AC Q51506;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCV-2000 (Rel. 40, Last annotation update)
DE REDOX-SENSITIVE TRANSCRIPTIONAL ACTIVATOR SOXR.
GN SOXR OR PA2273.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=97197508; PubMed=9045804;
RA Liao X., Hancock R.E.W.;
RT "Identification of a penicillin-binding protein 3 homolog, pbp3x, in
Pseudomonas aeruginosa: gene cloning and growth phase-dependent
expression."
RL J. Bacteriol. 179:1490-1496(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PAO1;
RX MEDLINE=20437337; PubMed=10984043;
RA Stever C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Gabler L.L., Goltzy L., Tolentino E., Westbrook-Medman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.S., Wu Z., Paulsen I.T.,
RA Reizer J., Salier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
opportunistic pathogen."
RL Nature 406:959-964(2000).
CC -!- FUNCTION: ACTIVATES THE TRANSCRIPTION OF THE SOXS GENE WHICH
ITSELF CONTROL THE SUPEROXIDE RESPONSE REGULATOR. SOXR CONTAINS A
2FE-2S IRON-SULFUR CLUSTER THAT MAY ACT AS A REDOX SENSOR SYSTEM
THAT RECOGNIZES SUPEROXIDE. THE VARIABLE REDOX STATE OF THE FE-S
CLUSTER MAY THUS BE EMPLOYED IN VIVO TO MODULATE THE
TRANSCRIPTIONAL ACTIVITY OF SOXR IN RESPONSE TO SPECIFIC TYPES OF
OXIDATIVE STRESS (BY SIMILARITY).
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE MERR FAMILY OF TRANSCRIPTIONAL
REGULATORS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X95517; CAA64771.1;
EMBL; AE004653; AAC05661.1;

```

DR InterPro: IPR000551; -
DR Pfam: PF00376; merr; 1.
DR PRINTS: PR00040; HTHMER.
DR PROSITE: PS00552; HTH_MERR_FAMILY; 1.
KM DNA-binding; Transcription regulation; Activator; Iron-sulfur.
FT DNA_BIND 12 31 H-H MOTIF (POTENTIAL).
FT DOMAIN 117 128 MIGHT BE PART OF A SENSOR REGION.
SQ SEQUENCE 156 AA; 16998 MW; 7B542A2380AD72C2 CRC64;

Query Match 46.7%; Score 57; DB 1; Length 156;
Best Local Similarity 76.9%; Pred. No. 0.16;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 7 GCIGPGCACLQAC 19
DB 116 GCIGCGCLSLQAC 128

RESULT 2
MT11_MYTED STANDARD; PRT; 72 AA.
AC P80246;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE METALLOTHIONEIN 10-I (MT-10-I).
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytilloida;
OC Mytilloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6550;
RN [1]
RP SEQUENCE.
RX MEDLINE=94062828; PubMed=8243463;
RA Mackay E.A., Overnell J., Dunbar B., Davidson I., Hunziker P.E.,
RA Kaegi J.H.R., Fothergill J.E.;
RT "Complete amino acid sequences of five dimeric and four monomeric
RT forms of metallothionein from the edible mussel Mytilus edulis.";
RL Eur. J. Biochem. 218:183-194(1993).
CC -I- FUNCTION: THE METALLOTHIONEINIS ARE INVOLVED IN THE CELLULAR
CC SEQUESTRATION OF TOXIC METAL IONS.
CC -I- SUBUNIT: MONOMER.
CC -I- INDUCTION: BY CADMIUM.
CC -I- SIMILARITY: BELONGS TO FAMILY 2 IN METALLOTHIONEIN SUPERFAMILY.
DR PIR: S39416; S39416.
DR HSSP: P15358; 1SKZ.
DR InterPro: IPR001008; -
DR Pfam: PF00131; metalthio; 1.
DR PRINTS: PR00875; MTMOLUSC.
KM Metal-binding; Metal-thiolate cluster; Chelation; Cadmium.
SQ SEQUENCE 72 AA; 7105 MW; 3EA9D959AE8B3B2 CRC64;

Query Match 41.0%; Score 50; DB 1; Length 72;
Best Local Similarity 61.5%; Pred. No. 0.87;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 7 GCIGPGCACLQAC 19
DB 17 GCIGCGCGCGDAC 29

RESULT 3
MT12_MYTED STANDARD; PRT; 72 AA.
AC P80247;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE METALLOTHIONEIN 10-II (MT-10-II).
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytilloida;

OC Mytilloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6550;
RN [1]

RP SEQUENCE.
RX MEDLINE=94062828; PubMed=8243463;
RA Mackay E.A., Overnell J., Dunbar B., Davidson I., Hunziker P.E.,
RA Kaegi J.H.R., Fothergill J.E.;
RT "Complete amino acid sequences of five dimeric and four monomeric
RT forms of metallothionein from the edible mussel Mytilus edulis.";
RL Eur. J. Biochem. 218:183-194(1993).
CC -I- FUNCTION: THE METALLOTHIONEINIS ARE INVOLVED IN THE CELLULAR
CC SEQUESTRATION OF TOXIC METAL IONS.
CC -I- SUBUNIT: MONOMER.
CC -I- INDUCTION: BY CADMIUM.
CC -I- SIMILARITY: BELONGS TO FAMILY 2 IN METALLOTHIONEIN SUPERFAMILY.
DR PIR: S39417; S39417.
DR HSSP: P15358; 1SKZ.
DR InterPro: IPR001008; -
DR Pfam: PF00131; metalthio; 1.
DR PRINTS: PR00875; MTMOLUSC.
KM Metal-binding; Metal-thiolate cluster; Chelation; Cadmium.
SQ SEQUENCE 72 AA; 7036 MW; 7FA9D959AE8B269 CRC64;

Query Match 41.0%; Score 50; DB 1; Length 72;
Best Local Similarity 61.5%; Pred. No. 0.87;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 7 GCIGPGCACLQAC 19
DB 17 GCIGCGCGCGDAC 29

RESULT 4
MT13_MYTED STANDARD; PRT; 72 AA.
AC P80248;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE METALLOTHIONEIN 10-III (MT-10-III).
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytilloida;
OC Mytilloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6550;
RN [1]
RP SEQUENCE.
RX MEDLINE=94062828; PubMed=8243463;
RA Mackay E.A., Overnell J., Dunbar B., Davidson I., Hunziker P.E.,
RA Kaegi J.H.R., Fothergill J.E.;
RT "Complete amino acid sequences of five dimeric and four monomeric
RT forms of metallothionein from the edible mussel Mytilus edulis.";
RL Eur. J. Biochem. 218:183-194(1993).
CC -I- FUNCTION: THE METALLOTHIONEINIS ARE INVOLVED IN THE CELLULAR
CC SEQUESTRATION OF TOXIC METAL IONS.
CC -I- SUBUNIT: MONOMER.
CC -I- INDUCTION: BY CADMIUM.
CC -I- SIMILARITY: BELONGS TO FAMILY 2 IN METALLOTHIONEIN SUPERFAMILY.
DR PIR: S39418; S39418.
DR HSSP: P04355; 4MT2.
DR InterPro: IPR001008; -
DR Pfam: PF00131; metalthio; 1.
DR PRINTS: PR00875; MTMOLUSC.
KM Metal-binding; Metal-thiolate cluster; Chelation; Cadmium.
SQ SEQUENCE 72 AA; 6949 MW; 648AF4576E80726E CRC64;

Query Match 41.0%; Score 50; DB 1; Length 72;
Best Local Similarity 61.5%; Pred. No. 0.87;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

OY 7 GCIGPCACIAC 19
DB 17 GCSGECRCGDAC 29

RESULT 5
MT14_MYTED STANDARD: PRT: 72 AA.
AC P80249:
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last annotation update)
DE METALLOTHIONEIN 10-1V (MF-10-1V).
OS Mytilus edulis (Blue mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloida;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6550;
RN [1]
RM
RX MEDLINE=94062828; PubMed=8243463;
RA Mackay A.A., Overnell J., Dunbar B., Davidson I., Hunziker P.E.,
RA Kaegi J.H.R., Fothergill J.E.;
RT "Complete amino acid sequences of five dimeric and four monomeric
RL forms of metallothionein from the edible mussel Mytilus edulis.";
RL Eur. J. Biochem. 218:183-194(1993).
CC -I- FUNCTION: THE METALLOTHIONEINS ARE INVOLVED IN THE CELLULAR
CC SEQUESTRATION OF TOXIC METAL IONS.
CC -I- SUBUNIT: MONOMER.
CC -I- INDUCTION: BY CADMIUM.
CC -I- SIMILARITY: BELONGS TO FAMILY 2 IN METALLOTHIONEIN SUPERFAMILY.
DR HSSP; P04355; 4MT2.
DR InterPro; IPR001008; -.
DR InterPro; IPR003019; -.
DR Pfam; PF00331; metalthio. 1.
DR PRINTS; PR00875; MTFMOLISC.
KW Metal-binding; Metal-thiolate cluster; Chelation; Cadmium.
SQ SEQUENCE 72 AA: 6979 MW: 7FA028637D837269 CRC64;

Query Match 41.0%; Score 50; DB 1; Length 72;
Best Local Similarity 61.5%; Pred. No. 0.87;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 7 GCIGPCACIAC 19
DB 17 GCSGECRCGDAC 29

RESULT 6
C79A_MOUSE STANDARD: PRT: 220 AA.
ID C79A_MOUSE
AC P11911:
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE B-CELL ANTIGEN RECEPTOR COMPLEX ASSOCIATED PROTEIN ALPHA-CHAIN
DE PRECURSOR (IG-ALPHA) (MB-1 MEMBRANE GLYCOPROTEIN) (SUPRACE-IGH-
DE ASSOCIATED PROTEIN) (MEMBRANE-BOUND IMMUNOGLOBULIN ASSOCIATED
DE PROTEIN) (CD79A).
GN CD79A OR IGA OR MB-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RM
RX MEDLINE=89091088; PubMed=2463161;
RA Sakaguchi N., Kashiwamura S., Kimoto M., Thalmann P., Melchers F.;
RT "B lymphocyte lineage-restricted expression of mb-1, a gene with CD3-
RL like structural properties.";
RL EMBO J. 7:3457-3464(1988).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BAB/C; TISSUE=Liver;
RX MEDLINE=90293481; PubMed=2358676;
RA Kashiwamura S.-I., Koyama T., Steinmetz M., Kimoto M.,
RA Sakaguchi N.;
RT "Structure of the murine mb-1 gene encoding a putative
RT sigm-associated molecule.";
RL J. Immunol. 145:337-343(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92347937; PubMed=1639443;
RA Flaszinkel H., Reth M.;
RT "Molecular cloning of the Ig-alpha subunit of the human B-cell
RT antigen receptor complex.";
RL Immunogenetics 36:266-269(1992).
CC -I- FUNCTION: ASSOCIATED TO SURFACE IGM-RECEPTOR; MAY BE INVOLVED IN
CC SIGNAL TRANSDUCTION.
CC -I- SUBUNIT: HETERODIMER OF ALPHA AND BETA CHAINS, DISULFIDE-LINKED.
CC -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -I- TISSUE SPECIFICITY: B-CELLS.
CC -I- PTM: PHOSPHORYLATED ON TYROSINE AS A RESULT OF B-CELL ACTIVATION.
CC -I- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X13450; CA31801.1; -.
DR EMBL; M31773; AAA39494.1; -.
DR PIR; S01648; S01648.
DR PIR; A43540; A43540.
DR MGD; MGI:101774; Cd79a.
KW Immunoglobulin domain; Transmembrane; Glycoprotein; B-cell; Signal;
KW Phosphorylation.
FT SIGNAL 1 22
FT CHAIN 23 220
FT FT 23 220
FT DOMAIN 23 137
FT TRANSSEM 138 159
FT DOMAIN 160 220
FT DOMAIN 43 108
FT DISULFID 50 101
FT DISULFID 113 113
FT CARBOHYD 58 58
FT CARBOHYD 68 68
FT CONFLICT 95 100
SQ SEQUENCE 220 AA: 24261 MW: A654123C58177B29 CRC64;

Query Match 41.0%; Score 50; DB 1; Length 220;
Best Local Similarity 43.8%; Pred. No. 2.1;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 VYESMCTIGPCACIQ 17
DB 15 LFLSYACIGPCQALR 30

RESULT 7
TXP3_APTSC STANDARD: PRT: 37 AA.
ID TXP3_APTSC
AC P49268;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE APTOTOXIN III (PARALYTIC PEPTIDE III) (PP III).
OS Aptostichus schlingeri (Trap-door spider).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;

```

CC Mygalomorphae; Cyrtachenidae; Aptostichus.
 OX NCBI_TaxID=12944;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Venom;
 RX MEDLINE=93069259; PubMed=1440641;
 RA Skinner W.S., Dennis P.A., Li J.P., Oulstad G.B.;
 RT "Identification of insecticidal peptides from venom of the trap-door
 spider, Aptostichus schlingeri (Ctenizidae).";
 RL Toxicon 30:1043-1050(1992).
 CC -1- FUNCTION: IS BOTH PARALYTIC AND LETHAL, WHEN INJECTED INTO
 CC LEPIDOPTERAN LARVAE. IS A SLOWER ACTING TOXIN, BEING LETHAL AT 24
 CC HR, BUT NOT PARALYTIC AT 1 HR POST-INJECTION.
 CC -1- PTM: FOUR DISULFIDE BONDS ARE PRESENT.
 CC -1- MISCELLANEOUS: LD(50) IS 0.50 MG/KG BY SUBCUTANEOUS INJECTION.
 CC -1- SIMILARITY: TO APTOTOXIN VII.
 KW Venom; Toxin.
 SQ SEQUENCE 37 AA; 3769 MW; CBD01091694E1908 CRC64;

Query Match 40.2%; Score 49; DB 1; Length 37;
 Best Local Similarity 58.3%; Pred. No. 0.7;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 3 YESMGCIGPCA 14
 Db 21 YNWMNCIGGCS 32

RESULT 8
 VC47_HSV11 STANDARD; PRT; 518 AA.
 AC 000139;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE 56 KDA SERINE PROTEASE (EC 3.4.21.-).
 GN 47.
 OS Ictalurid herpesvirus 1 (Channel catfish virus) (CCV).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC unclassified Herpesviridae.
 OX NCBI_TaxID=10401;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AUBURN 1;
 RA MEDLINE=92087490; PubMed=1727613;
 RX Davison A.J.;
 RT "Channel catfish virus: a new type of herpesvirus.";
 RL Virology 186:9-14(1992).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
 CC SUBTILASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M75136; AAA88150.1; -
 DR PIR: C36791; PRBET1.
 DR InterPro: IPR000209; -
 DR InterPro: IPR002884; -
 DR Pfam: PF01483; P; 1.
 DR PROSITE: PS00136; SUBTILASE_ASP; FALSE_NEG.
 DR PROSITE: PS00137; SUBTILASE_HIS; FALSE_NEG.
 DR PROSITE: PS00138; SUBTILASE_SER; FALSE_NEG.
 KW Hydrolase; Serine protease.
 FT ACT_SITE 334 334
 SQ SEQUENCE 518 AA; 55984 MW; DAAAF919A975460D CRC64;

Query Match 39.3%; Score 48; DB 1; Length 518;
 Best Local Similarity 61.5%; Pred. No. 8.2;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 YVESMGCIGPCA 14
 Db 232 VSESMGCVDDGAA 244

RESULT 9
 RRPO_SBMV STANDARD; PRT; 956 AA.
 ID RRPO_SBMV
 AC P21405.
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE PROBABLE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48).
 OS Southern bean mosaic virus (SBMV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Sobemovirus.
 OX NCBI_TaxID=12139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COMPEA;
 RX MEDLINE=88044510; PubMed=2823471;
 RA Mu S., Rinehart C.A., Kaesberg P.;
 RT "Sequence and organization of southern bean mosaic virus genomic
 RT RNA.";
 RL Virology 161:73-80(1987).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M23021; AAA46565.1; -
 DR PIR: B33739; RRBWSC.
 DR InterPro: IPR001795; -
 DR Pfam: PF02123; Luteo_ORF3; 1.
 DR PRINTS: PR00914; LVTRUSRNAPOL.
 KW Transferase; RNA-directed RNA polymerase.
 KW SEQUENCE 956 AA; 104844 MW; ABB961C9B76082D8 CRC64;

Query Match 39.3%; Score 48; DB 1; Length 956;
 Best Local Similarity 57.1%; Pred. No. 13;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 QVESMGCIGPCA 14
 Db 937 QAVPSMGJHGPCS 950

RESULT 10
 YEYH_ECOLI STANDARD; PRT; 385 AA.
 ID YEYH_ECOLI
 AC P33361; P76435;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE HYPOTHETICAL ABC TRANSPORTER PERMEASE PROTEIN YEYH.
 GN YEYH.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / BHB2600;
 RA Richterich P., Lakey N., Gryan G., Jaehn L., Mintz L., Robison K.,
 RA Church G.M.;

RL Submitted (OCT-1993) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blotner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Valdes J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 CC -1- FUNCTION: PROBABLY PART OF A BINDING-PROTEIN-DEPENDENT TRANSPORT
 CC SYSTEM. PROBABLY RESPONSIBLE FOR THE TRANSLLOCATION OF THE
 CC SUBSTRATE ACROSS THE MEMBRANE.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
 CC (POTENTIAL).
 CC -1- SIMILARITY: WITH: INTEGRAL MEMBRANE COMPONENTS OF OTHER BINDING-
 CC PROTEIN-DEPENDENT TRANSPORT SYSTEMS. BELONGS TO THE CSTM
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: U00007; AAA60493.1; ALT_INIT.
 CC DR EMBL: AE000302; AAC75191.1; -.
 CC DR Ecogene: EG12011; yehY.
 CC DR InterPro: IPR000515; -.
 CC DR Pfam: PF00528; BPD_transp.1.
 CC DR PROSITE: PS00402; BPD_TRANSP_INN_MEMBR.1.
 CC KW Hypothetical protein; Transport; Transmembrane; Inner membrane.
 FT TRANSMEM 6 26
 FT TRANSMEM 45 65 POTENTIAL.
 FT TRANSMEM 70 90 POTENTIAL.
 FT TRANSMEM 107 127 POTENTIAL.
 FT TRANSMEM 181 201 POTENTIAL.
 FT TRANSMEM 225 245 POTENTIAL.
 FT TRANSMEM 248 268 POTENTIAL.
 FT TRANSMEM 323 343 POTENTIAL.
 FT TRANSMEM 347 367 POTENTIAL.
 FT CONFLICT 346 385 ILSSAIDLVLTGVPIVAVLTDAFLDLLLALKVKRND
 FT -> AAKORH (IN REF. 1).
 SQ SEQUENCE 385 AA; 41138 MW; E57055E3A2B141CC CRC64;
 Query Match 38.5%; Score 47; DB 1; Length 385;
 Best Local Similarity 57.1%; Pred. No. 8.9;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 Oy 6 WCGICPGCACLOAC 19
 Db 50 WMLGVGCAMLTAC 63
 RESULT 11
 PROP_CAVPO STANDARD: PRT; 470 AA.
 ID PROP_CAVPO
 AC 064181;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PROPERDIN PRECURSOR.
 GN PFC.
 OS *Cavia porcellus* (Guinea pig).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Hysticognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=Spleen;
 RX MEDLINE=96148617; PubMed=8550088;
 RA Maves K.K., Guenther S.T., Densen P., Moser D.R., Weiler J.M.;
 RT Cloning and characterization of the cDNA encoding guinea-pig
 RT properdin: a comparison of properdin from three species.*;
 RL Immunology 86:475-479(1995).
 CC -1- FUNCTION: A POSITIVE REGULATOR OF THE ALTERNATE PATHWAY OF
 CC COMPLEMENT. IT BINDS TO AND STABILIZES THE C3- AND C5-CONVERTASE
 CC ENZYME COMPLEXES.
 CC -1- SIMILARITY: CONTAINS 6 TSP TYPE-1 DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: S81116; AAB35918.1; -.
 CC DR InterPro: IPR000884; -.
 CC DR Pfam: PF00090; tsp_1; 6.
 CC DR PROSITE: PS50092; TSP1; 6.
 CC KW Signal; Complement alternate pathway; Glycoprotein; Repeat.
 FT SIGNAL 1 26
 FT CHAIN 27 470
 FT DOMAIN 76 133 TSP TYPE-1 1.
 FT DOMAIN 134 190 TSP TYPE-1 2.
 FT DOMAIN 191 254 TSP TYPE-1 3.
 FT DOMAIN 255 312 TSP TYPE-1 4.
 FT DOMAIN 313 376 TSP TYPE-1 5.
 FT DOMAIN 379 438 TSP TYPE-1 6.
 FT CARBOHYD 429 429 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 470 AA; 51431 MW; FDC2B393DC7EC15F CRC64;
 Query Match 37.7%; Score 46; DB 1; Length 470;
 Best Local Similarity 46.7%; Pred. No. 15;
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 Oy 5 SWGICPGCACLOAC 19
 Db 194 AWGPGWSSCLSSC 208
 RESULT 12
 CYGR_ARBPV STANDARD: PRT; 986 AA.
 ID CYGR_ARBPV
 AC P11528;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE RESACT RECEPTOR PRECURSOR (GUANYLATE CYCLASE) (EC 4.6.1.2).
 OS *Arbacia punctulata* (Punctate sea urchin).
 CC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 CC Echinoidea; Euechinoidea; Echinacea; Arbacoidea; Arbaciidae; Arbacia.
 OX NCBI_TaxID=7641;
 RN [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RC TISSUE=Testis;
 RX MEDLINE=88318927; PubMed=2901039;
 RA Singh S., Lowe D.G., Thorpe D.S., Rodriguez H., Kuang W.-J.,
 RA Dangolt L.J., Chinkers M., Goedel D.V., Garbers D.L.;
 RT "Membrane guanylate cyclase is a cell-surface receptor with homology
 RT to protein kinases.";
 RL Nature 334:708-712(1988).
 CC -1- FUNCTION: IMPLICATED AS A CELL-SURFACE RECEPTOR ON SPERMATOZOA
 CC FOR 'RESACT' A CHEMOTACTIC PEPTIDE, AND ON VARIOUS OTHER CELLS
 CC AS A RECEPTOR FOR ATRIAL Natriuretic Peptide.
 CC -1- CATALYTIC ACTIVITY: GTP = 3,'5'-CYCLIC GMP + PYROPHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- SIMILARITY: SOME SIMILARITY WITH CONSERVED REGION OF CATALYTIC
 CC DOMAIN OF PROTEIN KINASES.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X12874; CAA31367.1; -.
 DR PIR: S05480; OYURCA.
 DR InterPro: IPR000719; -.
 DR InterPro: IPR001828; -.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR Pfam: PF00069; PKinase; 2.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase;
 KW GMP synthesis; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 986 RESACT RECEPTOR.
 FT DOMAIN 22 507 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 508 528 POTENTIAL.
 FT DOMAIN 529 986 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 568 836 PROTEIN KINASE LIKE.
 FT CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 361 361 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 410 410 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 986 AA: 111284 MW: B40238A74CCAF52 CRC64;

Query Match 37.7%; Score 46; DB 1; Length 986;
 Best Local Similarity 77.8%; Pred. No. 26;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 7 GCIGPGCAC 15
 I I I I I I I
 Db 112 GVIGPGCGC 120

RESULT 13
 CYS-STRPU STANDARD; PRT; 1125 AA.
 ID CYS-STRPU STANDARD; PRT; 1125 AA.
 AC P16065;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE SPERACT RECEPTOR PRECURSOR (GUANYLATE CYCLASE) (BC 4.6.1.2).
 OS Strongylocentrotus purpuratus (Purple sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinidea; Echinoidea; Echinozoa; Strongylocentrotidae;
 OC Strongylocentrotus.
 OX NCBI_TaxId=7668;
 RN NCBI_TaxId=7668;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89197965; PubMed=2564849;
 RA Thorpe D.S., Garbers D.L.;
 RT "The membrane form of guanylate cyclase. Homology with a subunit of
 RT the cytoplasmic form of the enzyme.";
 RL J. Biol. Chem. 264:6545-6549(1989).
 CC -1- FUNCTION: IMPLICATED AS A CELL-SURFACE RECEPTOR ON SPERMATOZOA
 CC FOR 'SPERACT' A CHEMOTACTIC PEPTIDE. AND ON VARIOUS OTHER CELLS
 CC AS A RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE.
 CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-CYCLIC GMP + PYROPHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- SIMILARITY: SOME SIMILARITY WITH CONSERVED REGION OF CATALYTIC
 CC DOMAIN OF PROTEIN KINASES.
 CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M22444; AAA30051.1; -.
 DR PIR: A33535; OYURCP.
 DR PIR: A30856; A30856.
 DR HSSP: Q02846; IAWL.
 DR InterPro: IPR000719; -.
 DR InterPro: IPR001054; -.
 DR InterPro: IPR001828; -.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR Pfam: PF00069; PKinase; 1.
 DR PROSITE: PS00452; GUANYLATE_CYCLASES_1; 1.
 DR PROSITE: PS0125; GUANYLATE_CYCLASES_2; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase;
 KW GMP synthesis; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 1125 SPERACT RECEPTOR.
 FT DOMAIN 22 510 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 511 531 POTENTIAL.
 FT DOMAIN 532 1125 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 571 839 PROTEIN KINASE LIKE.
 FT CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 409 409 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1125 AA: 126256 MW: EAF6757BCF92782 CRC64;

Query Match 37.7%; Score 46; DB 1; Length 1125;
 Best Local Similarity 77.8%; Pred. No. 29;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 7 GCIGPGCAC 15
 I I I I I I I
 Db 112 GVIGPGCGC 120

RESULT 14
 MT2A-RABIT STANDARD; PRT; 62 AA.
 ID MT2A-RABIT STANDARD; PRT; 62 AA.
 AC P18055;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE METALLOTHIONEIN-11A (MT-2A).
 OS METALLOTHIONEIN-11A (MT-2A).
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxId=9986;
 RN NCBI_TaxId=9986;
 RP SEQUENCE.
 RC STRAIN=NEW ZEALAND WHITE; TISSUE=Liver, and Kidney;
 RX MEDLINE=92140112; PubMed=1779803;
 RA Hunziker P.E.;
 RT "Amino acid sequence determination.";
 RL Meth. Enzymol. 205:421-426(1991).
 RN [2]
 RP SEQUENCE.
 RC STRAIN=NEW ZEALAND WHITE; TISSUE=Liver, and Kidney;
 RX MEDLINE=95169065; PubMed=7864820;
 RA Hunziker P.E., Kaur P., Wan M., Kaenzig A.;
 RT "Primary structures of seven metallothioneins from rabbit tissue.";
 RL Biochem. J. 306:265-270(1995).
 RN [3]
 RP STRUCTURE BY NMR. AND SEQUENCE.
 RX MEDLINE=86220204; PubMed=3709538;
 RA Wagner G., Neuhaus D., Worgatter E., Vasak M., Kaegi J.H.R.,
 RA Wuerlich K.;
 RT "Sequence-specific 1H-NMR assignments in rabbit-liver
 RT metallothionein-2.";
 RL Eur. J. Biochem. 157:275-289(1986).
 RN [4]

RP STRUCTURE BY NMR.
RX MEDLINE=86171767; PubMed=3959079;
RA Wagner G., Neuhau D., Worgotter E., Vasak M., Kaegi J.H.R.,
Wuetrich K.;
RT "Nuclear magnetic resonance identification of 'half-turn' and 3(10)-
helix secondary structure in rabbit liver metallothionein-2.";
RL J. Mol. Biol. 187:131-135(1986).
RN [5]
RP STRUCTURE BY NMR.
RX MEDLINE=88332988; PubMed=3418714;
RA Arseniev A., Schultze P., Worgotter E., Braun W., Wagner G., Vasak M.,
Kaegi J.H.R., Wuetrich K.;
RT "Three-dimensional structure of rabbit liver [Cd]metallothionein-2a
in aqueous solution determined by nuclear magnetic resonance.";
RL J. Mol. Biol. 201:637-657(1988).
CC -1- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
RESIDUES THAT BIND VARIOUS HEAVY METALS; THESE PROTEINS ARE
TRANSCRIPTIONALLY REGULATED BY BOTH HEAVY METALS AND
GLUCOCORTICOIDS.
CC -1- SUBUNIT: MONOMER.
CC -1- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:
FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA
DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11
CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE
BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 IN METALLOTHIONEIN SUPERFAMILY.
DR PDB: 1MBR; 15-JUL-92.
DR PDB: 2MRB; 15-JUL-92.
DR InterPro: IPR000006; -;
DR InterPro: IPR003019; -;
DR Pfam: PF00131; metalthio. 1.
DR PRINTS: PR00860; MTVERTERATE.
DR PROSITE: PS00203; METALLOTHIONEIN_VRT; 1.
KW Metal-binding; Metal-thiolate cluster; Chelation; Zinc; Copper;
KW Cadmium; Acetylation; 3D-structure.
FT MOD_RES 1 1 ACETYLATION.
FT DOMAIN 1 30 BETA.
FT METAL 31 62 ALPHA.
FT METAL 5 5 CLUSTER B.
FT METAL 7 7 CLUSTER B.
FT METAL 14 14 CLUSTER B.
FT METAL 16 16 CLUSTER B.
FT METAL 20 20 CLUSTER B.
FT METAL 22 22 CLUSTER B.
FT METAL 25 25 CLUSTER B.
FT METAL 27 27 CLUSTER B.
FT METAL 30 30 CLUSTER B.
FT METAL 34 34 CLUSTER A.
FT METAL 35 35 CLUSTER A.
FT METAL 37 37 CLUSTER A.
FT METAL 38 38 CLUSTER A.
FT METAL 42 42 CLUSTER A.
FT METAL 45 45 CLUSTER A.
FT METAL 49 49 CLUSTER A.
FT METAL 51 51 CLUSTER A.
FT METAL 58 58 CLUSTER A.
FT METAL 60 60 CLUSTER A.
FT METAL 61 61 CLUSTER A.
FT TURN 28 29
SQ SEQUENCE 62 AA; 6083 MW; C3FEA7B701081B3E CRC64;

Query Match 37.3%; Score 45.5; DB 1; Length 62;
Best Local Similarity 57.1%; Pred. No. 3.3;
Matches 8; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

OY 8 CIGPCA-CLOACL 20
1 1111 1 1 1
Db 37 CCPCGCAKCAQCGI 50

RESULT 15
MT2C_RABIT

ID MT2C_RABIT STANDARD: PRT: 62 AA.
AC P80290;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE METALLOTHIONEIN-TIC (MT-2C).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE.
RC STRAIN=NEW ZEALAND WHITE; TISSUE=Liver, and Kidney;
RX MEDLINE=95169065; PubMed=7864620;
RA Hunziker P.E., Kaur P., Wan M., Kaenzig A.;
RT "Primary structures of seven metallothioneins from rabbit tissue.";
RL Biochem. J. 306:265-270(1995).
CC -1- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
RESIDUES THAT BIND VARIOUS HEAVY METALS; THESE PROTEINS ARE
TRANSCRIPTIONALLY REGULATED BY BOTH HEAVY METALS AND
GLUCOCORTICOIDS.
CC -1- SUBUNIT: MONOMER.
CC -1- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:
FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA
DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11
CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE
BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 IN METALLOTHIONEIN SUPERFAMILY.
DR HSSP: P18055; 1MBR.
DR InterPro: IPR000006; -;
DR InterPro: IPR003019; -;
DR Pfam: PF00131; metalthio. 1.
DR PRINTS: PR00860; MTVERTERATE.
DR PROSITE: PS00203; METALLOTHIONEIN_VRT; 1.
KW Metal-binding; Metal-thiolate cluster; Chelation; Zinc; Copper;
KW Cadmium; Acetylation.
FT MOD_RES 1 1 ACETYLATION.
FT DOMAIN 1 30 BETA.
FT METAL 31 62 ALPHA.
FT METAL 5 5 CLUSTER B.
FT METAL 7 7 CLUSTER B.
FT METAL 14 14 CLUSTER B.
FT METAL 16 16 CLUSTER B.
FT METAL 20 20 CLUSTER B.
FT METAL 22 22 CLUSTER B.
FT METAL 25 25 CLUSTER B.
FT METAL 27 27 CLUSTER B.
FT METAL 30 30 CLUSTER B.
FT METAL 34 34 CLUSTER A.
FT METAL 35 35 CLUSTER A.
FT METAL 37 37 CLUSTER A.
FT METAL 38 38 CLUSTER A.
FT METAL 42 42 CLUSTER A.
FT METAL 45 45 CLUSTER A.
FT METAL 49 49 CLUSTER A.
FT METAL 51 51 CLUSTER A.
FT METAL 58 58 CLUSTER A.
FT METAL 60 60 CLUSTER A.
FT METAL 61 61 CLUSTER A.
FT TURN 28 29
SQ SEQUENCE 62 AA; 6113 MW; D6EBA7B70112AB3E CRC64;

Query Match 37.3%; Score 45.5; DB 1; Length 62;
Best Local Similarity 57.1%; Pred. No. 3.3;
Matches 8; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

OY 8 CIGPCA-CLOACL 20
1 1111 1 1 1
Db 37 CCPCGCAKCAQCGI 50

Job time: 87 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 15:44:57 ; Search time 33.25 Seconds

(without alignments)
79.582 Million cell updates/sec

Title: US-09-609-721-14

Perfect score: 122

Sequence: 1 OYESWGCIGECACACAC 20

Scoring table: BLOSUM62

Searched: 425026 segs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SPITREMBL_16:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_unclassified:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	51	41.8	148	2	09KN81	09KN81 vibrio chol
2	50	41.0	73	5	062554	062554 mytilus edu
3	50	41.0	73	5	062555	062555 mytilus edu
4	50	41.0	73	5	09U569	09U569 perna virid
5	50	41.0	73	5	09U568	09U568 perna virid
6	50	41.0	75	5	096388	096388 perna virid
7	50	41.0	659	1	028894	028894 archaeglob
8	50	41.0	709	3	09P3D2	09P3D2 neurospora
9	49	40.2	351	14	09E1H5	09E1H5 meleagrid h
10	49	40.2	351	14	09PPR9	09PPR9 meleagrid h
11	49	40.2	1063	13	09PMG1	09PMG1 oryzias lat
12	48	39.3	67	2	09JX24	09JX24 neisseria m
13	48	39.3	71	2	09JYX6	09JYX6 neisseria m
14	47.5	38.9	161	14	066618	066618 equine herp
15	46.5	38.1	315	5	093532	093532 caenorhabdi
16	46	37.7	230	10	09LK02	09LK02 arabidopsis
17	46	37.7	231	10	09FNH6	09FNH6 arabidopsis
18	46	37.7	411	3	074310	074310 schizosacch
19	46	37.7	673	3	013650	013650 schizosacch

20	46	37.7	1125	5	027669	027669 hemientrot
21	46	37.7	1127	5	09UBP5	09UBP5 diadema set
22	45.5	37.3	62	6	09TUI5	09TUI5 canis famli
23	45.5	37.3	387	2	09KG75	09KG75 bacillus ha
24	45.5	37.3	731	5	09N8D2	09N8D2 trypanosoma
25	45	36.9	175	2	09S255	09S255 streptomyce
26	45	36.9	248	3	043033	043033 schizosacch
27	45	36.9	398	8	034939	034939 leishmania
28	45	36.9	1136	11	004692	004692 mus musculu
29	44.5	36.5	47	5	09N9H0	09N9H0 venerupis p
30	44.5	36.5	57	5	09N9H2	09N9H2 ruditapes p
31	44.5	36.5	59	5	09N9H1	09N9H1 ruditapes d
32	44.5	36.5	75	5	09N1N5	09N1N5 crassostrea
33	44.5	36.5	107	5	09NG19	09NG19 crassostrea
34	44.5	36.5	137	14	082466	082466 unidentified
35	44.5	36.5	180	1	030080	030080 archaeglob
36	44.5	36.5	214	1	093725	093725 pyrobaculum
37	44.5	36.5	831	5	017516	017516 caenorhabdi
38	44.5	36.5	2824	13	09W7R3	09W7R3 brachydanio
39	44	36.1	128	5	044567	044567 caenorhabdi
40	44	36.1	348	5	09VVW4	09VVW4 drosophila
41	44	36.1	433	10	09SYC4	09SYC4 arabidopsis
42	44	36.1	449	5	09NKA5	09NKA5 drosophila
43	44	36.1	790	10	091Q24	091Q24 arabidopsis
44	44	36.1	1055	13	09YCW3	09YCW3 oryzias lat
45	43.5	35.7	113	10	082765	082765 chlamydomon

ALIGNMENTS

RESULT 1
ID 09KN81 PRELIMINARY: PRT: 148 AA.
AC 09KN81:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE SOXR PROTEIN.
GN VCA0084.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_taxid=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utherback T., Fleischman R.D., Niernan W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
CC -!- SIMILARITY: BELONGS TO THE MERR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC EMBL: AE004351; AAF95998.1; -
DR TIGR: VCA0084; -
DR InterPro: IPR000551; -
DR Pfam: PF00376; merr.1.
DR PRINTS: PR00040; HTMMERR.
DR PROSITE: PS00552; HTM_MERR_FAMILY; 1.
KW SMART: SM00422; HTM_MERR; 1.
KW DNA-binding; Transcription regulation.
SQ SEQUENCE 148 AA; 16986 MW; 6987166A8CD35508 CRC64;

Query Match 41.8%; Score 51; DB 2; Length 148;
Best Local Similarity 61.5%; Pred. No. 3.5;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 7 GCIGPGCACLQAC 19
 |||| 11 1::1
 Db 109 GCIGGCGCLSLESC 121

RESULT 2
 ID 062554 PRELIMINARY; PRT; 73 AA.
 AC 062554
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE METALLOTHIONEIN 10 IB.
 OS Mytilus edulis (Blue mussel).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Mytilus.
 OX NCBI_TaxID=6550;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=DIGESTIVE GLAND;
 RA Barsyte D., White K.N., Lovejoy D.A.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ005452; CAA06549.1; -
 DR InterPro: IPR000561; -
 DR InterPro: IPR001008; -
 DR InterPro: IPR003019; -
 DR Pfam: PF00131; metalthio; 1.
 DR PRINTS: PR00875; MTMOLLUSC.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 SQ SEQUENCE 73 AA; 7220 MW; 76797FB0FCD7B3B2 CRC64;

Query Match 41.0%; Score 50; DB 5; Length 73;
 Best Local Similarity 61.5%; Pred. No. 2.7;
 Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 GCIGPGCACLQAC 19
 |||| 11 1::1
 Db 18 GCSEGGCRCGDAC 30

RESULT 3
 ID 062555 PRELIMINARY; PRT; 73 AA.
 AC 062555
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE METALLOTHIONEIN 10 II.
 OS Mytilus edulis (Blue mussel).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Mytilus.
 OX NCBI_TaxID=6550;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=DIGESTIVE GLAND;
 RA Barsyte D., White K.N., Lovejoy D.A.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ005453; CAA06550.1; -
 DR InterPro: IPR000561; -
 DR InterPro: IPR001008; -
 DR InterPro: IPR003019; -
 DR Pfam: PF00131; metalthio; 1.
 DR PRINTS: PR00875; MTMOLLUSC.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 SQ SEQUENCE 73 AA; 7153 MW; 20CEFA4CB3A0CDE19 CRC64;

Query Match 41.0%; Score 50; DB 5; Length 73;
 Best Local Similarity 61.5%; Pred. No. 2.7;
 Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 GCIGPGCACLQAC 19
 |||| 11 1::1
 Db 18 GCSEGGCRCGDAC 30

RESULT 4
 ID 090569 PRELIMINARY; PRT; 73 AA.
 AC 090569
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE METALLOTHIONEIN 1.
 OS Perna viridis.
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Perna.
 OX NCBI_TaxID=73031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=99379847; PubMed=10451422;
 RA Khoo H.W., Patel K.H.;
 RT "Metallothionein cDNA, promoter, and genomic sequences of the tropical
 green mussel, Perna viridis."
 RL J. Exp. Zool. 284:445-453(1999).
 DR EMBL: AF092971; AAF22486.1; -
 DR InterPro: IPR000561; -
 DR InterPro: IPR001008; -
 DR InterPro: IPR003019; -
 DR Pfam: PF00131; metalthio; 1.
 DR PRINTS: PR00875; MTMOLLUSC.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 SQ SEQUENCE 73 AA; 7254 MW; 42293A9547DF5CA8 CRC64;

Query Match 41.0%; Score 50; DB 5; Length 73;
 Best Local Similarity 61.5%; Pred. No. 2.7;
 Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 GCIGPGCACLQAC 19
 |||| 11 1::1
 Db 18 GCSEGGCRCGDAC 30

RESULT 5
 ID 090568 PRELIMINARY; PRT; 73 AA.
 AC 090568
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE METALLOTHIONEIN 2.
 OS Perna viridis.
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Perna.
 OX NCBI_TaxID=73031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=99379847; PubMed=10451422;
 RA Khoo H.W., Patel K.H.;
 RT "Metallothionein cDNA, promoter, and genomic sequences of the tropical
 green mussel, Perna viridis."
 RL J. Exp. Zool. 284:445-453(1999).
 DR EMBL: AF092972; AAF22487.1; -
 DR InterPro: IPR000561; -
 DR InterPro: IPR001008; -
 DR InterPro: IPR003019; -
 DR Pfam: PF00131; metalthio; 1.
 DR PRINTS: PR00875; MTMOLLUSC.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 SQ SEQUENCE 73 AA; 7277 MW; ABCED9547DF5CA7 CRC64;

Query Match 41.0%; Score 50; DB 5; Length 73;

GN UL16.
 OS Meleagrid herpesvirus 1 (herpesvirus of turkeys).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gamaherpesvirinae.
 OX NCBI_TaxID=37108;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FC126;
 RA Kingham B.J., Zelink V., Kopacek J., Majerciak V., Ney E., Chen Y.,
 RA Schmidt C.J.;
 RT *Coding potential of herpesvirus of turkey: comparative genetic
 RT analysis of Marek's disease serologically related viruses.*;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF282130; AAG30055.1; -
 SO SEQUENCE 351 AA; 39083 MW; 1D527D3846EE9E82 CRC64;

Query Match 40.2%; Score 49; DB 14; Length 351;
 Best Local Similarity 44.4%; Pred. No. 14;
 Matches 8; Conservative 2; Mismatches 6; Indels 2; Gaps 1;
 QY 5 SMGC--IGPGCACLQACL 20
 |||| :| | | | :
 Db 244 SMGCGTVGHACICKGPCM 261

RESULT 10
 ID Q9DPR9 PRELIMINARY; PRT; 351 AA.
 AC Q9DPR9;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE UL16 TEGUMENT PROTEIN.
 GN HT023.
 OS Meleagrid herpesvirus 1 (herpesvirus of turkeys).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gamaherpesvirinae.
 OX NCBI_TaxID=37108;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FC126;
 RA Afonso C.L., Tulman E.R., Lu Z., Zsak L., Rock D.L., Kutish G.K.;
 RT "The genome of turkey herpesvirus.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF291866; AAG45753.1; -
 SO SEQUENCE 351 AA; 39137 MW; FD81745E4ADA713E CRC64;

Query Match 40.2%; Score 49; DB 14; Length 351;
 Best Local Similarity 44.4%; Pred. No. 14;
 Matches 8; Conservative 2; Mismatches 6; Indels 2; Gaps 1;
 QY 5 SMGC--IGPGCACLQACL 20
 |||| :| | | | :
 Db 244 SMGCGTVGHACICKGPCM 261

RESULT 11
 ID Q9PMG1 PRELIMINARY; PRT; 1063 AA.
 AC Q9PMG1;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE MEMBRANE GUANYLYL CYCLASE.
 GN OIGC2.
 OS Oryzias latipes (Medaka fish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
 OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
 OX NCBI_TaxID=8090;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Yamagami S., Muramatsu R., Suzuki N.;
 RT "Oryzias latipes mRNA for membrane guanylyl cyclase OIGC2, complete.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB030274; BAA62623.1; -
 DR HSSP: Q02846; 1AML.
 DR InterPro: IPR000719; -
 DR InterPro: IPR001054; -
 DR InterPro: IPR001170; -
 DR InterPro: IPR001245; -
 DR InterPro: IPR001828; -
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF00211; guanylate_cyc; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00255; NATPEPTIDER.
 DR PRINTS: PR00109; TYRKINASE.
 DR PROSITE: PS00452; GUANYLATE_CYCLASES; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR SMART: SM00044; CYCC; 1.
 KW ATP-binding; lyase; transferase; Tyrosine-protein kinase.
 SO SEQUENCE 1063 AA; 120364 MW; D3E672FEC6362313 CRC64;

Query Match 40.2%; Score 49; DB 13; Length 1063;
 Best Local Similarity 58.3%; Pred. No. 37;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 3 YESMGCIGPGCA 14
 | : | ||||| :
 Db 118 YKPMAFISPGCS 129

RESULT 12
 ID Q9JXZ4 PRELIMINARY; PRT; 67 AA.
 AC Q9JXZ4;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE HYPOTHETICAL PROTEIN NMB1825.
 GN NMB1825.
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MC58 / SEROGROUP B;
 RX MEDLINE=20175755; PubMed=10710307;
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Nelson W.C., Gwin M.L., DeBoy R., Peterson J.D., Hickey E.R.,
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
 RA Cotton M.D., Uitterback T.R., Khouli H., Qin H., Vamathevan J.,
 RA Gill J., Scarlato V., Maignani V., Pizza M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
 RT MC58.";
 RL Science 287:1809-1815(2000).
 DR EMBL: AE002532; AAF42160.1; -
 DR TIGR: NMB1825; -
 KW Hypothetical protein.
 SO SEQUENCE 67 AA; 7569 MW; A0F482D38A0DEC86 CRC64;

Query Match 39.3%; Score 48; DB 2; Length 67;
 Best Local Similarity 56.2%; Pred. No. 4.8;
 Matches 9; Conservative 1; Mismatches 4; Indels 2; Gaps 1;
 QY 4 ESMGCIGPGC--ACIQ 17
 ||| | | | | :
 Db 25 ESMCGSDCGEACIO 40


```

RESULT 13
Q9JYX6 PRELIMINARY: PRT; 71 AA.
AC Q9JYX6:
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE HYPOTHETICAL PROTEIN NMA0634.
GN NMA0634.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Felwell T., Hamlin N., Holtroyd S.,
RA Jagers K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis 22491."
RL Nature 404:502-506(2000).
DR EMBL: AL162753; CAB83924.1;
KW - Hypothetical protein.
SQ SEQUENCE 71 AA; 8027 MW; 275966B1E012D125 CRC64;

```

```

Query Match          39.3%; Score 48; DB 2; Length 71;
Best Local Similarity 56.2%; Pred. No. 5.1;
Matches 9; Conservative 1; Mismatches 4; Indels 2; Gaps 1;
QY 4 ESMGICGPGC-ACIQ 17
Db 27 ESMGICGPGCDAQIQ 42

```

```

RESULT 14
Q66618 PRELIMINARY: PRT; 161 AA.
AC Q66618:
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DE 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE ORF 12.
OS Equine herpesvirus type 2 (strain 86/87).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae.
OX NCBI_TaxID=82831;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95302501; PubMed=783207;
RA Telford E.A., Watson M.S., Aird H.C., Perry J., Davison A.J.;
RT "The DNA sequence of equine herpesvirus 2."
RL J. Mol. Biol. 249:520-528(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Telford E.A.R.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: U20824; AAC13801.1;
SQ SEQUENCE 161 AA; 17750 MW; C49163D3E254F773 CRC64;

```

```

Query Match          38.9%; Score 47.5; DB 14; Length 161;
Best Local Similarity 46.7%; Pred. No. 12;
Matches 7; Conservative 2; Mismatches 3; Indels 3; Gaps 1;
QY 5 SMCICGPGCAGC 19
Db 1 SMCICGPGCAGC 19

```

Db 140 SWGC---SCCFKSC 151

```

RESULT 15
Q93532 PRELIMINARY: PRT; 315 AA.
AC Q93532:
DT 01-FEB-1997 (TREMblrel. 02, Created)
DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)
DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE F20D1.8 PROTEIN.
GN F20D1.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea;
OC Rhabdilitida; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Burton J.;
RT Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten N., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Woldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
DR EMBL: Z78542; CAB01742.1;
DR InterPro: IPR001507;
DR SMART: SM00241; 2P; 1.
SQ SEQUENCE 315 AA; 35785 MW; FEDF94021B39CF00 CRC64;

```

```

Query Match          38.1%; Score 46.5; DB 5; Length 315;
Best Local Similarity 32.0%; Pred. No. 30;
Matches 8; Conservative 2; Mismatches 8; Indels 7; Gaps 1;
QY 2 VYEMSCGPGC-----CACTAC 19
Db 232 VYEMSCGPGCNOTSPCYTVHSC 256

```

Search completed: November 21, 2001, 15:47:34
 Job time: 157 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 15:44:57 ; Search time 32.74 Seconds

(without alignments)
37.034 Million cell updates/sec

Title: US-09-609-721-14

Perfect score: 122

Sequence: 1 QVIESMGCIGPGCACLQACL 20

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0601:*

- 1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
- 2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
- 3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
- 4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
- 5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
- 6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
- 7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
- 8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
- 9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
- 10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
- 11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
- 12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
- 13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
- 14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
- 15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
- 16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
- 17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
- 18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
- 19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
- 20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
- 21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
- 22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	122	100.0	20	22	ERB2 binding pept
2	122	100.0	20	22	Amino acid sequenc
3	122	100.0	30	22	HER2 peptide ligand
4	122	100.0	42	22	HER2 peptide ligand
5	122	100.0	45	22	HER2 peptide ligand
6	122	100.0	46	22	Amino acid sequenc
7	109	89.3	19	22	HER2 peptide ligand
8	95	77.9	20	22	ERB2 binding pept
9	94	77.0	17	22	ERB2 binding pept
10	90	73.8	17	22	ERB2 binding pept
11	89	73.0	20	22	ERB2 binding pept

12	88	72.1	17	22	AAB76373	ERB2 binding pept
13	88	72.1	20	22	AAB76367	ERB2 binding pept
14	87	71.3	17	22	AAB76372	ERB2 binding pept
15	87	71.3	20	22	AAB76360	ERB2 binding pept
16	87	71.3	20	22	AAB76364	ERB2 binding pept
17	87	71.3	20	22	AAB76366	ERB2 binding pept
18	87	71.3	20	22	AAB76376	ERB2 binding pept
19	86	70.5	17	22	AAB76370	ERB2 binding pept
20	86	70.5	17	22	AAB76374	ERB2 binding pept
21	86	70.5	20	22	AAB76359	ERB2 binding pept
22	86	70.5	17	22	AAB76362	ERB2 binding pept
23	85	69.7	17	22	AAB76369	ERB2 binding pept
24	85	69.7	20	22	AAB76361	ERB2 binding pept
25	85	69.7	20	22	AAB76436	HER2 binding pept
26	84	68.9	17	22	AAB76368	ERB2 binding pept
27	74	60.7	20	22	AAB76408	ERB2 binding pept
28	74	60.7	20	22	AAB76414	ERB2 binding pept
29	72	59.0	20	22	AAB76411	ERB2 binding pept
30	72	59.0	20	22	AAB76418	ERB2 binding pept
31	72	59.0	20	22	AAB76433	HER2 binding pept
32	69	56.6	17	22	AAB76377	HER2 binding pept
33	69	56.6	17	22	AAB76438	HER2 binding pept
34	68	55.7	16	22	AAB76392	ERB2 binding pept
35	68	55.7	20	22	AAB76379	ERB2 binding pept
36	68	55.7	20	22	AAB76417	HER2 binding pept
37	68	55.7	20	22	AAB76434	HER2 binding pept
38	67	54.9	16	22	AAB76383	ERB2 binding pept
39	67	54.9	16	22	AAB76386	ERB2 binding pept
40	67	54.9	16	22	AAB76390	ERB2 binding pept
41	67	54.9	20	22	AAB76350	ERB2 binding pept
42	67	54.9	20	22	AAB76393	ERB2 binding pept
43	67	54.9	20	22	AAB76412	ERB2 binding pept
44	67	54.9	20	22	AAB76435	HER2 binding pept
45	66	54.1	16	22	AAB76385	ERB2 binding pept

ALIGNMENTS

RESULT 1

ID AAB76363 standard; Peptide: 20 AA.

AC AAB76363:

DT 10-APR-2001 (first entry)

XX

DE ErbB2 binding peptide amino acid sequence SEQ ID 14.

XX

XX Human; erbB2; HER2; cancer; nervous system disease; stroke; ischaemia;

KW metabolic disorder; nutritional deficiency; Alzheimer's disease;

KW Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;

KW Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.

XX

OS Synthetic.

XX

PN MO200101748-A2.

PD 11-JAN-2001.

XX

XX 30-JUN-2000; 2000WO-US18283.

PF 02-JUL-1999; 99US-0142232.

XX

XX (GETH) GENENTECH INC.

PA

XX

PI Dennis MS;

XX

DR WPI; 2001-123046/13.

XX

XX Non-naturally occurring peptide ligands which compete for binding human

PT erbB2 gene products; useful for treating e.g. Alzheimer's disease;

PT multiple sclerosis and diabetic neuropathy -

XX Disclosure; Figure 16; 116pp; English.
PS
XX
CC This invention relates to non-naturally occurring peptide ligands which
CC bind to the human erbB2 gene product ErbB2 (also known as HER2). Peptides
CC represented in AAB/6350 - AAB/6420 and AAB/6432 - AAB/6509 are examples
CC of the ErbB2 binding ligands of the invention. Sequences
CC AAB/6421 - AAB/6431 represent antibody Fc amino acid sequences used in
CC the isolation of the peptides of the invention. The peptides compete for
CC binding ErbB2 with naturally occurring ligands, and may be used to treat
CC disorders characterized by over expression of HER2/ErbB2 such as cancers,
CC diseases of the nervous system, musculature and epithelia, e.g. nervous
CC system damage resulting from trauma, surgery, strokes, ischaemia,
CC infection, metabolic disorders, nutritional deficiency or toxic agents.
CC In particular the synthetic peptide ligands may be used to treat
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis,
CC Huntington's chorea, Down's syndrome, nerve deafness, Meniere's disease
CC and neuropathy associated with diabetes.
CC
XX
SQ Sequence 20 AA;

Query Match 100.0%; Score 122; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QVYESMGCTGPGCACTQACL 20
Db 1 qyewsgctgpgcacclgac1 20
|||||

RESULT 2
AAB/6430
ID AAB/6430 standard; Protein; 20 AA.
XX
AC AAB/6430;
XX
DT 10-APR-2001 (first entry)
XX
DE Amino acid sequence of 1.1FT-2 SEQ ID 81.
XX-
XX Human; erbB2; HER2; cancer; nervous system disease; stroke; ischaemia;
XX metabolic disorder; nutritional deficiency; Alzheimer's disease;
XX Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
XX Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.
XX
OS Synthetic.
XX
XX WO200101748-A2.
XX
XX 11-JAN-2001.
XX
XX 30-JUN-2000; 2000WO-US18283.
XX
XX 02-JUL-1999; 99US-0142232.
XX
XX (GETH) GENENTECH INC.
XX
XX Dennis MS;
XX
XX WPI: 2001-123048/13.
XX
XX
XX Non-naturally occurring peptide ligands which compete for binding human
XX ErbB2 gene products, useful for treating e.g. Alzheimer's disease,
XX multiple sclerosis and diabetic neuropathy -
XX
XX
XX Disclosure; Figure 8A; 116pp; English.
XX
XX This invention relates to non-naturally occurring peptide ligands which
XX bind to the human erbB2 gene product ErbB2 (also known as HER2). Peptides
XX represented in AAB/6350 - AAB/6420 and AAB/6432 - AAB/6509 are examples
XX of the ErbB2 binding ligands of the invention. Sequences
XX AAB/6421 - AAB/6431 represent antibody Fc amino acid sequences used in

CC the isolation of the peptides of the invention. The peptides compete for
CC binding ErbB2 with naturally occurring ligands, and may be used to treat
CC disorders characterized by over expression of HER2/ErbB2 such as cancers,
CC diseases of the nervous system, musculature and epithelia, e.g. nervous
CC system damage resulting from trauma, surgery, strokes, ischaemia,
CC infection, metabolic disorders, nutritional deficiency or toxic agents.
CC In particular the synthetic peptide ligands may be used to treat
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis,
CC Huntington's chorea, Down's syndrome, nerve deafness, Meniere's disease
CC and neuropathy associated with diabetes.
CC
XX
SQ Sequence 20 AA;

Query Match 100.0%; Score 122; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QVYESMGCTGPGCACTQACL 20
Db 1 qyewsgctgpgcacclgac1 20
|||||

RESULT 3
AAB/67215
ID AAB/67215 standard; peptide; 30 AA.
XX
AC AAB/67215;
XX
DT 10-APR-2001 (first entry)
XX
DE HER2 peptide ligand #4.
XX
XX Fusion protein; immunoglobulin; multidimerization domain; ligand.
XX
XX Synthetic.
XX
XX WO200102440-A1.
XX
XX 11-JAN-2001.
XX
XX 30-JUN-2000; 2000WO-US18185.
XX
XX 02-JUL-1999; 99US-0142088.
XX
XX (GETH) GENENTECH INC.
XX
XX Dennis MS; Lazarus RA;
XX
XX WPI: 2001-123106/13.
XX
XX
XX Novel fusion polypeptides comprising a peptide ligand domain which
XX functions to target hybrid molecule to target cell, and immunoglobulin
XX constant region multidimerization domain -
XX
XX
XX Example 1; Page 39; 69pp; English.
XX
XX The present invention relates to a fusion protein, comprising a
XX peptide ligand and an immunoglobulin (Ig) constant region
XX multidimerization domain (ID). The hybrid molecules comprising
XX the peptide ligands and their functional derivatives can be used
XX in the same applications as, a peptide ligand can be used. For
XX example the peptide ligand can bind ErbB2. The peptide ligand
XX may bind to and inhibit the activity associated with a particular
XX target molecule.
XX
XX
SQ Sequence 30 AA;

Query Match 100.0%; Score 122; DB 22; Length 30;
Best Local Similarity 100.0%; Pred. No. 1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PT Non-naturally occurring peptide ligands which compete for binding human
PT Erb2 gene products, useful for treating e.g. Alzheimer's disease,
PS multiple sclerosis and diabetic neuropathy -
PS Disclosure: Figure 9A-C; 116pp; English.
XX
CC This invention relates to non-naturally occurring peptide ligands which
CC bind to the human erb2 gene product Erb2 (also known as HER2). Peptides
CC represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples
CC of the Erb2 binding ligands of the invention. Sequences
CC AAB76421 - AAB76431 represent antibody Fc amino acid sequences used in
CC the isolation of the peptides of the invention. The peptides compete for
CC binding Erb2 with naturally occurring ligands, and may be used to treat
CC disorders characterized by over expression of HER2/Erb2 such as cancers,
CC diseases of the nervous system, musculature and epithelia, e.g. nervous
CC system damage resulting from trauma, surgery, strokes, ischaemia,
CC infection, metabolic disorders, nutritional deficiency or toxic agents.
CC In particular the synthetic peptide ligands may be used to treat
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis,
CC Huntington's chorea, Down's syndrome, nerve deafness, Meniere's disease
CC and neuropathy associated with diabetes.
XX
SQ Sequence 46 AA:

Query Match 100.0%; Score 122; DB 22; Length 46;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVYESWGCIGPGCAGLQACL 20
Db 1 qvyeswgcigpgcagclqacl 20
|||||
1 qvyeswgcigpgcagclqacl 20

RESULT 7
AAB67212
ID AAB67212 standard; peptide: 19 AA.
XX
AC AAB67212;
XX
DT 10-APR-2001 (first entry)
XX
DE HER2 peptide ligand #1.
XX
KW Fusion protein; immunoglobulin; multidimerization domain; ligand.
XX
OS Synthetic.
XX
PN WO200102440-A1.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18185.
XX
PR 02-JUL-1999; 99US-0142088.
XX
PA (GETH) GENENTECH INC.
XX
PI Dennis MS; Lazarus RA;
XX
DR WPI; 2001-123106/13.
XX
PT Novel fusion polypeptides comprising a peptide ligand domain which
PT functions to target hybrid molecule to target cell, and immunoglobulin
PT constant region multimerization domain -
XX
PS Example 1; Page 39; 69pp; English.
XX
CC The present invention relates to a fusion protein, comprising a
CC peptide ligand and an immunoglobulin (Ig) constant region
CC multimerization domain (Ib). The hybrid molecules comprising
CC the peptide ligands and their functional derivatives can be used
CC in the same applications as, a peptide ligand can be used. For

CC example the peptide ligand can bind Erb2. The peptide ligand
CC may bind to and inhibit the activity associated with a particular
CC target molecule.
XX
SQ Sequence 19 AA:

Query Match 89.3%; Score 109; DB 22; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.7e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVYESWGCIGPGCAGLQ 18
Db 1 qvyeswgcigpgcagclq 18
|||||
1 qvyeswgcigpgcagclq 18

RESULT 8
AAB76358
ID AAB76358 standard; peptide: 20 AA.
XX
AC AAB76358;
XX
DT 10-APR-2001 (first entry)
XX
DE Erb2 binding peptide amino acid sequence SEQ ID 9.
XX
KW Human: erb2; HER2; cancer; nervous system disease; stroke; ischaemia;
KW metabolic disorder; nutritional deficiency; Alzheimer's disease;
KW Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
KW Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.
XX
OS Synthetic.
XX
PN WO200101748-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18283.
XX
PR 02-JUL-1999; 99US-0142232.
XX
PA (GETH) GENENTECH INC.
XX
PI Dennis MS;
XX
DR WPI; 2001-123048/13.
XX
PT Non-naturally occurring peptide ligands which compete for binding human
PT Erb2 gene products, useful for treating e.g. Alzheimer's disease,
PT multiple sclerosis and diabetic neuropathy -
XX
PS Disclosure: Figure 16; 116pp; English.
XX
CC This invention relates to non-naturally occurring peptide ligands which
CC bind to the human erb2 gene product Erb2 (also known as HER2). Peptides
CC represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples
CC of the Erb2 binding ligands of the invention. Sequences
CC AAB76421 - AAB76431 represent antibody Fc amino acid sequences used in
CC the isolation of the peptides of the invention. The peptides compete for
CC binding Erb2 with naturally occurring ligands, and may be used to treat
CC disorders characterized by over expression of HER2/Erb2 such as cancers,
CC diseases of the nervous system, musculature and epithelia, e.g. nervous
CC system damage resulting from trauma, surgery, strokes, ischaemia,
CC infection, metabolic disorders, nutritional deficiency or toxic agents.
CC In particular the synthetic peptide ligands may be used to treat
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis,
CC Huntington's chorea, Down's syndrome, nerve deafness, Meniere's disease
CC and neuropathy associated with diabetes.
XX
SQ Sequence 20 AA:

Query Match 77.9%; Score 95; DB 22; Length 20;

```
QY      4 ESWGCI GPGCACLQACL 20
        1:||||| 1 1 1 1
Db      1 eawgcigpgcgclmacl 17
```

QY	4	ESWGCIGPGCACLQACL	20
Db	1	epwgcigpgcacllgcl	17

DT 10-APR-2001 (first entry)

DE ErbB2 binding peptide amino acid sequence SEQ ID 16.
XX
XX Human: erbB2; HER2; cancer; nervous system disease; stroke; ischaemia;
KW metabolic disorder; nutritional deficiency; Alzheimer's disease;
KW Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
KW Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.
XX
OS Synthetic.
XX
PN WO200101748-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18283.
XX
PR 02-JUL-1999; 99US-0142232.
XX
PA (GETH) GENENTECH INC.
XX
PI Dennis MS;
XX
DR WPI; 2001-123048/13.
XX
PT Non-naturally occurring peptide ligands which compete for binding human
PT ErbB2 gene products, useful for treating e.g. Alzheimer's disease,
PT multiple sclerosis and diabetic neuropathy -
XX
PS Disclosure; Figure 16; 116pp; English.
XX
XX This invention relates to non-naturally occurring peptide ligands which
CC bind to the human erbB2 gene product ErbB2 (also known as HER2). Peptides
CC represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples
CC of the ErbB2 binding ligands of the invention. Sequences
CC AAB76421 - AAB76431 represent antibody Fc amino acid sequences used in
CC the isolation of the peptides of the invention. The peptides compete for
CC binding ErbB2 with naturally occurring ligands, and may be used to treat
CC disorders characterized by over expression of HER2/ErbB2 such as cancers,
CC diseases of the nervous system, musculature and epithelia, e.g. nervous
CC system damage resulting from trauma, surgery, strokes, ischaemia,
CC infection, metabolic disorders, nutritional deficiency or toxic agents.
CC In particular the synthetic peptide ligands may be used to treat
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis,
CC Huntington's chorea, Down's syndrome, nerve deafness, Meniere's disease
CC and neuropathy associated with diabetes.
XX
XX Sequence 20 AA:
SQ
Query Match 73.0%; Score 89; DB 22; Length 20;
Best Local Similarity 82.4%; Pred. No. 8.3e-05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 4 ESWGCTGPGCACLQACL 20
DB 4 eswgctgpgccclngcl 20
RESULT 12
AAB76373
ID AAB76373 standard; Peptide; 17 AA.
XX
AC AAB76373;
XX
DT 10-APR-2001 (first entry)
XX
DE ErbB2 binding peptide amino acid sequence SEQ ID 24.
XX
KW Human; erbB2; HER2; cancer; nervous system disease; stroke; ischaemia;
KW metabolic disorder; nutritional deficiency; Alzheimer's disease;
KW Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
KW Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.
XX
OS Synthetic.

XX
XX WO200101748-A2.
PN
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18283.
XX
PR 02-JUL-1999; 99US-0142232.
XX
PA (GETH) GENENTECH INC.
XX
PI Dennis MS;
XX
DR WPI; 2001-123048/13.
XX
PT Non-naturally occurring peptide ligands which compete for binding human
PT ErbB2 gene products, useful for treating e.g. Alzheimer's disease,
PT multiple sclerosis and diabetic neuropathy -
XX
PS Disclosure; Figure 16; 116pp; English.
XX
XX This invention relates to non-naturally occurring peptide ligands which
CC bind to the human erbB2 gene product ErbB2 (also known as HER2). Peptides
CC represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples
CC of the ErbB2 binding ligands of the invention. Sequences
CC AAB76421 - AAB76431 represent antibody Fc amino acid sequences used in
CC the isolation of the peptides of the invention. The peptides compete for
CC binding ErbB2 with naturally occurring ligands, and may be used to treat
CC disorders characterized by over expression of HER2/ErbB2 such as cancers,
CC diseases of the nervous system, musculature and epithelia, e.g. nervous
CC system damage resulting from trauma, surgery, strokes, ischaemia,
CC infection, metabolic disorders, nutritional deficiency or toxic agents.
CC In particular the synthetic peptide ligands may be used to treat
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis,
CC Huntington's chorea, Down's syndrome, nerve deafness, Meniere's disease
CC and neuropathy associated with diabetes.
XX
XX Sequence 17 AA:
SQ
Query Match 72.1%; Score 88; DB 22; Length 17;
Best Local Similarity 76.5%; Pred. No. 9.7e-05;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 4 ESWGCTGPGCACLQACL 20
DB 1 eswgctgpgckclmgcl 17
RESULT 13
AAB76367
ID AAB76367 standard; Peptide; 20 AA.
XX
AC AAB76367;
XX
DT 10-APR-2001 (first entry)
XX
DE ErbB2 binding peptide amino acid sequence SEQ ID 18.
XX
KW Human; erbB2; HER2; cancer; nervous system disease; stroke; ischaemia;
KW metabolic disorder; nutritional deficiency; Alzheimer's disease;
KW Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
KW Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.
XX
OS Synthetic.
XX
PN WO200101748-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18283.
XX
PR 02-JUL-1999; 99US-0142232.

XX (GETH) GENENTECH INC.
XX
XX
PI Dennis MS;
XX
DR WPI; 2001-123048/13.
XX
PT Non-naturally occurring peptide ligands which compete for binding human
PT erb2 gene products, useful for treating e.g. Alzheimer's disease,
PT multiple sclerosis and diabetic neuropathy -
XX
PS Disclosure: Figure 16; 116pp; English.
XX
CC This invention relates to non-naturally occurring peptide ligands which
CC bind to the human erb2 gene product ErbB2 (also known as HER2). Peptides
CC represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples
CC of the ErbB2 binding ligands of the invention. Sequences
CC AAB76421 - AAB76431 represent antibody Fc amino acid sequences used in
CC the isolation of the peptides of the invention. The peptides compete for
CC binding ErbB2 with naturally occurring ligands, and may be used to treat
CC disorders characterized by over expression of HER2/ErbB2 such as cancers,
CC diseases of the nervous system, musculature and epithelia, e.g. nervous
CC system damage resulting from trauma, surgery, strokes, ischaemia,
CC infection, metabolic disorders, nutritional deficiency or toxic agents.
CC In particular the synthetic peptide ligands may be used to treat
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis,
CC Huntington's chorea, Down's syndrome, nerve deafness, Meniere's disease
CC and neuropathy associated with diabetes.
XX
SQ Sequence 20 AA;

Query Match 72.1%; Score 88; DB 22; Length 20;
Best Local Similarity 66.7%; Pred. No. 0.00011;
Matches 12; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 3 YESMGLGPGCAGLQACL 20
: | ||||| | :
Db 3 hepgcigpgcaglnci 20

RESULT 14
AAB76372
ID AAB/6372 standard; Peptide; 17 AA.
XX
AC AAB76372;
XX
DT 10-APR-2001 (first entry)
XX
DE ErbB2 binding peptide amino acid sequence SEQ ID 23.
XX
KW Human; erbB2; HER2; cancer; nervous system disease; stroke; ischaemia;
KW metabolic disorder; nutritional deficiency; Alzheimer's disease;
KW Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
KW Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.
XX
OS Synthetic.
XX
PN WO200101748-A2.
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18283.
XX
PR 02-JUL-1999; 99US-0142232.
XX
PA (GETH) GENENTECH INC.
XX
PI Dennis MS;
XX
DR WPI; 2001-123048/13.
PT Non-naturally occurring peptide ligands which compete for binding human

PT erb2 gene products, useful for treating e.g. Alzheimer's disease,
PT multiple sclerosis and diabetic neuropathy -
XX
PS Disclosure: Figure 16; 116pp; English.
XX
CC This invention relates to non-naturally occurring peptide ligands which
CC bind to the human erb2 gene product ErbB2 (also known as HER2). Peptides
CC represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples
CC of the ErbB2 binding ligands of the invention. Sequences
CC AAB76421 - AAB76431 represent antibody Fc amino acid sequences used in
CC the isolation of the peptides of the invention. The peptides compete for
CC binding ErbB2 with naturally occurring ligands, and may be used to treat
CC disorders characterized by over expression of HER2/ErbB2 such as cancers,
CC diseases of the nervous system, musculature and epithelia, e.g. nervous
CC system damage resulting from trauma, surgery, strokes, ischaemia,
CC infection, metabolic disorders, nutritional deficiency or toxic agents.
CC In particular the synthetic peptide ligands may be used to treat
CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis,
CC Huntington's chorea, Down's syndrome, nerve deafness, Meniere's disease
CC and neuropathy associated with diabetes.
XX
SQ Sequence 17 AA;

Query Match 71.3%; Score 87; DB 22; Length 17;
Best Local Similarity 76.5%; Pred. No. 0.00013;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 ESMGLGPGCAGLQACL 20
: | ||||| | :
Db 1 eewgcigpgcaglnci 17

RESULT 15
AAB76360
ID AAB76360 standard; Peptide; 20 AA.
XX
AC AAB76360;
XX
DT 10-APR-2001 (first entry)
XX
DE ErbB2 binding peptide amino acid sequence SEQ ID 11.
XX
KW Human; erbB2; HER2; cancer; nervous system disease; stroke; ischaemia;
KW metabolic disorder; nutritional deficiency; Alzheimer's disease;
KW Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
KW Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.
XX
OS Synthetic.
XX
PN WO200101748-A2.
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18283.
XX
PR 02-JUL-1999; 99US-0142232.
XX
PA (GETH) GENENTECH INC.
XX
PI Dennis MS;
XX
DR WPI; 2001-123048/13.
XX
PT Non-naturally occurring peptide ligands which compete for binding human
PT erb2 gene products, useful for treating e.g. Alzheimer's disease,
PT multiple sclerosis and diabetic neuropathy -
XX
PS Disclosure: Figure 16; 116pp; English.
XX
CC This invention relates to non-naturally occurring peptide ligands which
CC bind to the human erb2 gene product ErbB2 (also known as HER2). Peptides
CC represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples

CC of the ErbB2 binding ligands of the invention. Sequences
 CC AAB76421 - AAB76431 represent antibody Fc amino acid sequences used in
 CC the isolation of the peptides of the invention. The peptides compete for
 CC binding ErbB2 with naturally occurring ligands, and may be used to treat
 CC disorders characterized by over expression of HER2/ErbB2 such as cancers,
 CC diseases of the nervous system, musculature and epithelia, e.g. nervous
 CC system damage resulting from trauma, surgery, strokes, ischemia,
 CC infection, metabolic disorders, nutritional deficiency or toxic agents.
 CC In particular the synthetic peptide ligands may be used to treat
 CC Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis,
 CC Huntington's chorea, Down's syndrome, nerve deafness, Meniere's disease
 CC and neuropathy associated with diabetes.

XX
 SQ Sequence 20 AA:

Query Match 71.3%; Score 87; DB 22; Length 20;
 Best Local Similarity 81.2%; Pred. No. 0.00015;
 Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 4 ESMGCIIPGCAICAC 19
 | | | | | | | | | | | | | | | | | | | | | |
 Db 4 edwgcigpgcacllxc 19

Search completed: November 21, 2001, 15:46:04
 Job time: 67 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2001, 15:44:57 ; Search time 19.72 Seconds

(Without alignments)
22.823 Million cell updates/sec

Title: US-09-609-721-14

Perfect score: 122
Sequence: 1 QVRESKCIGPCACIQAACL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	41.0	220	2	US-08-417-495-26
2	50	41.0	220	2	US-08-284-391B-26
3	50	41.0	220	4	US-09-218-950-26
4	50	41.0	220	5	PCT-US92-01785-26
5	50	41.0	220	5	PCT-US95-00454-26
6	49	40.2	243	4	US-09-413-814-20
7	48	39.3	19	2	US-08-751-305-33
8	48	39.3	2293	4	US-09-368-590-2
9	46	37.7	102	3	US-08-974-022-53
10	46	37.7	102	4	US-08-795-445A-53
11	46	37.7	102	4	US-08-795-447A-53
12	46	37.7	102	4	US-08-974-186-53
13	46	37.7	102	4	US-08-795-446B-53
14	45	36.9	624	3	US-08-422-108-1
15	45	36.9	782	2	US-09-146-283-4
16	45	36.9	782	3	US-08-579-823A-4
17	45	36.9	782	4	US-09-344-195-4
18	45	36.9	1255	1	US-08-467-083-68
19	45	36.9	1255	1	US-08-414-417B-68
20	45	36.9	1255	2	US-08-484-438-8
21	45	36.9	1255	2	US-08-486-348A-68
22	45	36.9	1255	2	US-08-625-101-2
23	45	36.9	1255	2	US-08-658-545B-68
24	45	36.9	1255	2	US-08-356-786-2
25	45	36.9	1255	3	US-08-466-680B-68
26	44.5	36.5	47	1	US-08-377-687-30
27	44.5	36.5	47	1	US-08-656-318A-12

28	44.5	36.5	47	2	US-08-777-192-30	Sequence 30, Appl
29	44.5	36.5	47	2	US-08-955-459-12	Sequence 12, Appl
30	44.5	36.5	47	4	US-08-971-962-30	Sequence 30, Appl
31	44	36.1	14	2	US-08-685-589A-189	Sequence 189, App
32	44	36.1	617	4	US-09-188-930-303	Sequence 303, App
33	43.5	35.7	52	1	US-07-609-716-68	Sequence 68, Appl
34	43.5	35.7	487	1	US-08-218-943-2	Sequence 2, Appl
35	43.5	35.7	1140	2	US-08-657-641-7	Sequence 7, Appl
36	43.5	35.7	1140	2	PCT-US94-07233-7	Sequence 7, Appl
37	43	35.2	1345	2	US-08-977-767-3	Sequence 3, Appl
38	42.5	34.8	34	1	US-08-321-668-7	Sequence 7, Appl
39	42.5	34.8	34	1	US-08-837-941-7	Sequence 7, Appl
40	42	34.4	38	2	US-08-903-516-47	Sequence 47, Appl
41	42	34.4	175	3	US-08-865-297-6	Sequence 6, Appl
42	42	34.4	581	3	US-09-221-235-8	Sequence 8, Appl
43	42	34.4	581	3	US-09-221-928-8	Sequence 8, Appl
44	42	34.4	581	4	US-09-221-527-8	Sequence 8, Appl
45	42	34.4	581	4	US-09-221-236-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-08-417-495-26
Sequence 26, Application US/08417495
Patent No. 5843728
GENERAL INFORMATION:
APPLICANT: Seed, Brian et al.
TITLE OF INVENTION: Redirection of Cellular Immunity by
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/417,495
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/203,866
FILING DATE:
APPLICATION NUMBER: US/07/847,566
FILING DATE:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 220 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: amino acids
US-08-417-495-26

Chimeras

Query Match 41.0%; Score 50; DB 2; Length 220;
Best Local Similarity 43.8%; Pred. No. 17;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 VYESWCGICGGCACIQ 17
DB 15 LFLSYACIGGCGCALR 30

RESULT 2

US-08-284-391B-26
Sequence 26, Application US/08284391B
Patent No. 5851828
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391B
FILING DATE: 02-AUG-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 220 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-284-391B-26

Query Match 41.0%; Score 50; DB 2; Length 220;
Best Local Similarity 43.8%; Pred. No. 17;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 VYESWCGICGGCACIQ 17
DB 15 LFLSYACIGGCGCALR 30

RESULT 3

US-09-218-950-26

Sequence 26, Application US/09218950
Patent No. 6284240
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/218,950
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391
FILING DATE: 02-AUG-1994
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 220 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-218-950-26

Query Match 41.0%; Score 50; DB 4; Length 220;
Best Local Similarity 43.8%; Pred. No. 17;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 VYESWCGICGGCACIQ 17
DB 15 LFLSYACIGGCGCALR 30

RESULT 4

PCT-US92-01785-26
Sequence 26, Application PC/TUS9201785
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor
TITLE OF INVENTION: Chimeras
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street

CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/01785
FILING DATE: 19920306
CLASSIFICATION: 530

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991

ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 220 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear

MOLECULE TYPE: amino acids

PCT-US92-01785-26

```

Query March 41.0% Score 50; DB 5; Length 220;
Best Local Similarity 43.8% Pred. NO. 17;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 VESWGCIGPGCACIQ 17
   :: 1:1:1111 1:
Db 15 LFLSYACIGPGCAQALR 30

RESULT 5
PCT-US95-00454 -26
; Sequence 26, Application PC/TUS9500454
GENERAL INFORMATION:
APPLICANT: Seed, Brian et al.
TITLE OF INVENTION: Targeted Cytolysis of HIV-Infected
TITLE OF INVENTION: Cells by Chimeric CD4 Receptor-
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00454
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/847,566
FILING DATE: March 6, 1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961

```

```

1      FILING DATE:  March 7, 1991
2      CLASSIFICATION:
3      ATTORNEY/AGENT INFORMATION:
4      NAME:  Clark, Paul T.
5      REGISTRATION NUMBER:  30,162
6      REFERENCE/DOCKET NUMBER:  00786/247001
7      TELECOMMUNICATION INFORMATION:
8      TELEPHONE:  (617) 542-5070
9      TELEFAX:  (617) 542-8906
10     TELEX:  200154
11     INFORMATION FOR SEQ ID NO:  26:
12     SEQUENCE CHARACTERISTICS:
13     LENGTH:  220 amino acids
14     TYPE:  amino acid
15     TOPOLOGY:  linear
16     MOLECULE TYPE:  amino acids
17
18     PCT-US95-00454-26

```

```

Query Match          41.0%; Score 50; DB 5; Length 220;
Best Local Similarity 43.8%; Pred. NO. 17;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY      2  VYESWGCIGPCACACLO 17
      : 1:1:1111 1:
Db      15  LFLSYACLGPGCQALR 30

RESULT 6
US-09-413-814-20
: Sequence 20, Application US/09413814
: Patent No. 6225064
: GENERAL INFORMATION:
: APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
: APPLICANT: Bristol-Myers Squibb, Co.
: APPLICANT: Bayer, Stefan
: APPLICANT: Bloecker, Helmut
: APPLICANT: Brandt, Petra
: APPLICANT: Cino, Paul M
: APPLICANT: Dougherty, Brian A
: APPLICANT: Goldberg, Steven L
: APPLICANT: Hoffer, Gerhard
: APPLICANT: Mueller, Joachim
: APPLICANT: Reichenbach, Hans
: TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polypeptide or
: FILE REFERENCE: PCT/US 99/23535
: CURRENT APPLICATION NUMBER: US/09/413,814
: CURRENT FILING DATE: 1999-10-07
: EARLIER APPLICATION NUMBER: DE 198 46 493.2
: NUMBER OF SEQ ID NOS: 107
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 20
: LENGTH: 243
: TYPE: PRT
: ORGANISM: Sorangium cellulosum
US-09-413-814-20

Query Match          40.2%; Score 49; DB 4; Length 243;
Best Local Similarity 42.1%; Pred. NO. 25;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY      1  QYESWGCIGPCACACLOAC 19
      : 1:1:1:111:1
Db      83  RVSVAWACLGMRACLETC 101

RESULT 7
US-08-751-305-33
: Sequence 33, Application US/08751305
: Patent No. 5965439

```

GENERAL INFORMATION:
APPLICANT: Tennet et al., Andrea J.
TITLE OF INVENTION: HOST DEFENSE ENHANCEMENT
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/751,305
FILING DATE: 18-NOV-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell, Jr., John R.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: 07306/012001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-751-305-33

Query Match 39.3%; Score 48; DB 2; Length 19;
Best Local Similarity 66.7%; Pred. No. 3.4;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 7 GCIGPGACACLOA 18
||| |||:
DB 4 GCIGATCACIGA 15

RESULT 8
US-09-368-590-2
Sequence 2, Application US/09368590
Patent No. 6187563
GENERAL INFORMATION:
APPLICANT: Solimena, Michele
TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
TITLE OF INVENTION: AUTOANTIGENS OF AUTOIMMUNE DISEASES
FILE REFERENCE: 101918-200 (OCR-941)
CURRENT APPLICATION NUMBER: US/09/368,590
CURRENT FILING DATE: 1999-08-04
EARLIER APPLICATION NUMBER: 60/095,657
EARLIER FILING DATE: 1998-08-07
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 2293
TYPE: PRT
ORGANISM: Human
US-09-368-590-2

Query Match 39.3%; Score 48; DB 4; Length 2293;
Best Local Similarity 64.3%; Pred. No. 2.4e+02;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 OYESMGCTGPGCA 14

Db 1350 KVRPSWCGGPGCA 1363
: | ||| ||| |

RESULT 9
US-08-974-022-53
Sequence 53, Application US/08974022
Patent No. 6015938
GENERAL INFORMATION:
APPLICANT: Boyle, William J.
APPLICANT: Lacey, David L.
APPLICANT: Calzone, Frank J.
APPLICANT: Chang, Ming-Shi
TITLE OF INVENTION: OSTEOPROTEGERIN
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,022
FILING DATE: 12-DEC-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/577,788
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-378
INFORMATION FOR SRO ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 102 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-974-022-53

Query Match 37.7%; Score 46; DB 3; Length 102;
Best Local Similarity 60.0%; Pred. No. 27;
Matches 9; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

OY 7 GCIGPG--CACLOAC 19
||| | ||| |
DB 46 GCCGGGTVACCTAC 60

RESULT 10
US-08-795-445A-53
Sequence 53, Application US/08795445A
Patent No. 6284485
GENERAL INFORMATION:
APPLICANT: Boyle, William J.
APPLICANT: Lacey, David L.
APPLICANT: Calzone, Frank J.
APPLICANT: Chang, Ming-Shi
TITLE OF INVENTION: OSTEOPROTEGERIN
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA


```

: NUMBER OF SEQUENCES: 53
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Amgen Inc.
: STREET: 1840 Dehavilland Drive
: CITY: Thousand Oaks
: STATE: California
: COUNTRY: USA
: ZIP: 91320-1789
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/795,446B
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/577,788
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Winter, Robert B.
: REFERENCE/DOCKET NUMBER: A-378
: INFORMATION FOR SEQ ID NO: 53:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 102 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-795-446B-53

```

```

Query Match      37.7%; Score 46; DB 4; Length 102;
Best Local Similarity 60.0%; Pred. No. 27;
Matches 9; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

```

```

QY 7 GCIGPG--CACIQAC 19
Db 46 GCCGGGTACACTAC 60

```

```

RESULT 14
US-08-422-108-1
: Sequence 1, Application US/08422108
: Patent No. 601567
: GENERAL INFORMATION:
: APPLICANT: Hudziak, Robert M.
: APPLICANT: Shepard, H. Michael
: APPLICANT: Ullrich, Axel
: TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 460 Point San Bruno Blvd
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: WinPatIn (Genentech)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/422,108
: FILING DATE: 14-Apr-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/355460
: FILING DATE: 13-DEC-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/048346

```

```

: FILING DATE: 15-APR-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/354319
: FILING DATE: 19-MAY-1989
: ATTORNEY/AGENT INFORMATION:
: NAME: Lee, Wendy M
: REGISTRATION NUMBER: 00,000
: REFERENCE/DOCKET NUMBER: 554C2D2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415/225-1994
: TELEFAX: 415/952-9881
: TELEEX: 910/371-7168
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 624 amino acids
: TYPE: Amino acid
: TOPOLOGY: Linear
: US-08-422-108-1

```

```

Query Match      36.9%; Score 45; DB 3; Length 624;
Best Local Similarity 58.3%; Pred. No. 1.8e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 8 CIGPCACIQAC 19
Db 483 CVGEGTACHQLC 494

```

```

RESULT 15
US-09-146-283-4
: Sequence 4, Application US/09146283
: Patent No. 5976546
: GENERAL INFORMATION:
: APPLICANT: Laus, Reiner
: APPLICANT: Ruegg, Curtis L.
: APPLICANT: Wu, Hongyu
: TITLE OF INVENTION: Immunostimulatory Compositions
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Denlinger & Associates
: STREET: 350 Cambridge Ave. Suite 250
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94306
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/146,283
: FILING DATE: 03-SEPT-1998
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: Judge, Linda R.
: REGISTRATION NUMBER: 42,702
: REFERENCE/DOCKET NUMBER: 7636-0010.21
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-324-0880
: TELEFAX: 650-324-0960
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 782 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHETICAL: NO
: ORIGINAL SOURCE:
: ORGANISM: homo sapiens
: INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
: US-09-146-283-4

```


Query Match 36.9%; Score 45; DB 2; Length 782;
Best Local Similarity 58.3%; Pred. No. 2.2e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Oy 8 CIGPGCAGCAGC 19
1:11111111
Db 504 CVGEGLAGCHQLC 515

Search completed: November 21, 2001, 15:45:25
Job time: 28 sec